

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2004, 19:57:07 ; Search time 197 Seconds
(without alignments)
1413.611 Million cell updates/sec

Title: US-10-078-650-12
Perfect score: 484
Sequence: 1 MDEGIPHLQERQLLEHRDFI.....PGNPESQAQEDPSQPGKEAP 484

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt 02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	440	90.9	482	1 BHB3_HUMAN	Q9C0J9 homo sapien
2	339	70.0	482	2 Q8TAT1	Q8tat1 homo sapien
3	114	23.6	513	2 Q7YQC9	Q7yqc9 canis famil
4	81	16.7	193	2 Q8CGH4	Q8cgh4 mus musculus
5	81	16.7	410	1 BHB3_MOUSE	Q99pv5 mus musculus
6	81	16.7	410	1 BHB3_RAT	Q35779 rattus norv
7	81	16.7	410	2 BAD18970	Bad18970 mus muscu
8	53	11.0	403	2 Q6NY50	Q6ny50 brachydanio
9	53	11.0	403	2 AAH66738	Aah66738 brachydan
10	53	11.0	409	2 Q6GNE9	Q6gne9 xenopus lae
11	52	10.7	411	1 BHB2_MOUSE	Q35185 mus musculus
12	52	10.7	411	1 BHB2_RAT	Q35780 rattus norv
13	52	10.7	411	2 Q76JQ4	Q76jq4 rattus norv
14	52	10.7	411	2 BAD01588	Bad01588 rattus no
15	52	10.7	411	2 BAC33281	Bac33281 mus muscu
16	52	10.7	411	2 BAC41056	Bac41056 mus muscu
17	52	10.7	412	1 BHB2_HUMAN	O14503 homo sapien
18	52	10.7	412	2 CAG33202	Cag33202 homo sapi
19	16	3.3	110	2 Q91LW0	Q91lw0 mus musculus
20	16	3.3	130	2 Q9TUC9	Q9tuc9 canis famil
21	16	3.3	131	2 Q9TST8	Q9tst8 felis silve
22	16	3.3	153	2 Q765Q0	Q765q0 rattus norv
23	16	3.3	153	2 BAD08305	Bad08305 rattus no
24	16	3.3	166	2 Q7PVF4	Q7pvf4 anopheles g
25	16	3.3	167	2 Q765P9	Q765p9 rattus norv
26	16	3.3	167	2 BAD08306	Bad08306 rattus no
27	16	3.3	213	2 Q9TUC8	Q9tuc8 monodelphis
28	16	3.3	286	2 Q6F5E0	Q6f5e0 mus musculus
29	16	3.3	304	1 GSH2_HUMAN	Q9bzm3 homo sapien
30	16	3.3	305	1 GSH2_MOUSE	P31316 mus musculus
31	16	3.3	314	1 PMXB_HUMAN	Q99453 homo sapien

32	16	3.3	314	1 PMXB_MOUSE	O35690 mus musculus
33	16	3.3	314	2 Q6PJD9	Q6pjd9 homo sapien
34	16	3.3	314	2 AAH17199	Aah17199 homo sapi
35	16	3.3	315	2 Q6YW21	Q6yw21 oryza sativ
36	16	3.3	315	2 BAD05688	Bad05688 oryza sat
37	16	3.3	315	2 BAD05845	Bad05845 oryza sat
38	16	3.3	331	1 SIA2_BRARE	Q7syl3 brachydanio
39	16	3.3	353	1 NK24_HUMAN	Q9h2z4 homo sapien
40	16	3.3	354	1 NK24_MOUSE	Q9egm3 mus musculus
41	16	3.3	362	1 NK23_MOUSE	P97334 mus musculus
42	16	3.3	364	1 NK23_HUMAN	Q8tau0 homo sapien
43	16	3.3	372	1 FBSh_HUMAN	Q9hah7 homo sapien
44	16	3.3	376	1 FXE1_HUMAN	O00358 homo sapien
45	16	3.3	378	1 FBSh_MOUSE	Q8r089 mus musculus

ALIGNMENTS

RESULT 1
BHB3_HUMAN
ID BHB3_HUMAN STANDARD; PRT; 482 AA.
AC Q9C0J9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (bHLHB3) (Differentially expressed in chondrocytes protein 2) (hDEC2) (Enhancer-of-split and hairy-related protein 1) (SHARP-1).
DE Name=BHLHB3; Synonyms=DEC2, SHARP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21092582; PubMed=11162494; DOI=10.1006/bbrc.2000.4133;
RA Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K., Yoshida E., Suardita K., Matsuda Y., Kato Y.;
RT "Molecular cloning and characterization of DEC2, a new member of basic helix-loop-helix proteins.";
RL Biochem. Biophys. Res. Commun. 280:164-171(2001).
RN [2]
RP FUNCTION.
RX MEDLINE=21226716; PubMed=11278948; DOI=10.1074/jbc.M011619200;
RA Garriga-Carut M., Roopra A., Buckley N.J.;
RT "The basic helix-loop-helix protein, SHARP-1, represses transcription by a histone deacetylase-dependent and histone deacetylase-independent mechanism.";
RL J. Biol. Chem. 276:14821-14828(2001).
CC -!- FUNCTION: May be a transcriptional repressor that represses both basal and activated transcription.
CC -!- SUBUNIT: Homodimerize.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in skeletal muscle and brain, moderately expressed in pancreas and heart, weakly expressed in placenta, lung, liver and kidney.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- SIMILARITY: Contains 1 orange domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB044088; BAB21502.1; -.
CC PIR; JC7583; JC7583.
CC Genew; HGNC:16617; BHLHB3.
CC MIM; 606200; -.
CC GO; GO:0005634; C:nucleus; NAS.

DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008283; P:cell proliferation; NAS.
DR GO; GO:0009887; P:organogenesis; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF07527; Hairy_orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Repressor; Transcription regulation.
FT DNA BIND 45 57 Basic motif.
FT DOMAIN 58 100 Helix-loop-helix motif.
FT DOMAIN 129 175 Orange.
FT DOMAIN 297 431 Ala/Gly-rich.
SQ SEQUENCE 482 AA; 50497 MW; 2BECDC2FDEB7CE14 CRC64;

Query Match 90.9%; Score 440; DB 1; Length 482;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
Db |||||
43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102

QY 105 TEQHQKIIALONGERSLSKSPIQSDLDADFHSQFQTCQVLSRFESWTREPVCVQL 164
Db |||||
103 TEQHQKIIALONGERSLSKSPIQSDLDADFHSQFQTCQVLSRFESWTREPVCVQL 162

QY 165 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 224
Db |||||
163 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 222

QY 225 RTQPSAEIAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 284
Db |||||
223 RTQPSAEIAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 282

QY 285 DSRGGSGGGPGGGGAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPA 344
Db |||||
283 DSRGGSGGGPGGGGAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPA 342

QY 345 AAAPFCPLPFCFLSPSAAAAYVQFLDKSGLEKYLPAAPAAAPFLLYPGIPAPAAAAA 404
Db |||||
343 AAAPFCPLPFCFLSPSAAAAYVQFLDKSGLEKYLPAAPAAAPFLLYPGIPAPAAAAA 402

QY 405 AAAAAAFAFPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLPFAGPRE 464
Db |||||
403 AAAAAAFAFPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLPFAGPRE 462

QY 465 PGNPESSAQEDPSQPGKEAP 484
Db |||||
463 PGNPESSAQEDPSQPGKEAP 482

RESULT 2
Q8TAT1 PRELIMINARY; PRT; 482 AA.
ID Q8TAT1
AC Q8TAT1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Basic helix-loop-helix domain containing, class B, 3.
GN Name=BHLHB3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa;
OC Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025968; AAH25968.1; -
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy_orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
SQ SEQUENCE 482 AA; 50525 MW; 23EBD420C6B0CE03 CRC64;

Query Match 70.0%; Score 339; DB 2; Length 482;
Best Local Similarity 99.8%; Pred. No. 3.8e-262;
Matches 439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
Db |||||
43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102

QY 105 TEQHQKIIALONGERSLSKSPIQSDLDADFHSQFQTCQVLSRFESWTREPVCVQL 164
Db |||||
103 TEQHQKIIALONGERSLSKSPIQSDLDADFHSQFQTCQVLSRFESWTREPVCVQL 162

QY 165 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 224
Db |||||
163 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 222

QY 225 RTQPSAEIAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 284
Db |||||
223 RTQPSAEIAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 282

QY 285 DSRGGSGGGPGGGGAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPA 344
Db |||||
283 DSRGGSGGGPGGGGAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPA 342

QY 345 AAAPFCPLPFCFLSPSAAAAYVQFLDKSGLEKYLPAAPAAAPFLLYPGIPAPAAAAA 404
Db |||||
343 AAAPFCPLPFCFLSPSAAAAYVQFLDKSGLEKYLPAAPAAAPFLLYPGIPAPAAAAA 402

QY 405 AAAAAAFAFPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLPFAGPRE 464
Db |||||
403 AAAAAAFAFPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLPFAGPRE 462

QY 465 PGNPESSAQEDPSQPGKEAP 484
Db |||||
463 PGNPESSAQEDPSQPGKEAP 482


```
RESULT 3
Q7YQC9
ID Q7YQC9 PRELIMINARY; PRT; 513 AA.
AC Q7YQC9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SHARP1 protein.
GN Name=SHARP1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=22791403; PubMed=12909371;
RA Kukekova A.V., Aguirre G.D., Acland G.M.;
RT "Cloning and characterization of canine SHARP1 and its evaluation as a
RT positional candidate for canine early retinal degeneration (erd).";
RL Gene 312:335-343(2003).
DR EMBL; AY204567; AAP12464.1; --
DR EMBL; AY204568; AAP12465.1; --
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
SQ SEQUENCE 513 AA; 52942 MW; CA7FC23C2E03DCBA CRC64;

Query Match 23.6%; Score 114; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 2.1e-82;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTCLKHLKALTAL 104
Db 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTCLKHLKALTAL 102

QY 105 TEQHQKIIALQNGERSLKSPQISDLDAFHSGFQTCAKEVLQYLSRFESWTPRE 158
Db 103 TEQHQKIIALQNGERSLKSPQISDLDAFHSGFQTCAKEVLQYLSRFESWTPRE 156

RESULT 4
Q8CGH4
ID Q8CGH4 PRELIMINARY; PRT; 193 AA.
AC Q8CGH4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bhlhb3 protein (Fragment).
GN Name=Bhlhb3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
```

```
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnersch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC037742; AAH37742.1; --
DR MGD; MGI:1930704; Bhlhb3.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
FT NON TER 193
SQ SEQUENCE 193 AA; 21929 MW; 854890594639A422 CRC64;

Query Match 16.7%; Score 81; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.2e-56;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTCLKHLKALTAL 104
Db 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTCLKHLKALTAL 102

QY 105 TEQHQKIIALQNGERSLKSP 125
Db 103 TEQHQKIIALQNGERSLKSP 123

RESULT 5
BHB3 MOUSE
ID BHB3 MOUSE STANDARD; PRT; 410 AA.
AC Q99PV5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (bHLHB3) (Differentially
DE expressed in chondrocytes protein 2) (mDEC2).
GN Name=Bhlhb3; Synonyms=Dec2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21092582; PubMed=1162494; DOI=10.1006/bbrc.2000.4133;
RA Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K.,
RA Yoshida E., Suardita K., Matsuda Y., Kato Y.;
RT "Molecular cloning and characterization of DEC2, a new member of basic
RT helix-loop-helix proteins.";
RL Biochem. Biophys. Res. Commun. 280:164-171(2001).
CC -!- FUNCTION: May be a transcriptional repressor that represses both
CC basal and activated transcription.
CC -!- SUBUNIT: Homodimerize.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in skeletal muscle, brain and lung.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- SIMILARITY: Contains 1 orange domain.
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB044090; BAB21503.1; --
PIR; JC7584; JC7584.
MGD; MGI:1930704; Bhlhb3.
InterPro; IPR001092; HLH_basic.
InterPro; IPR003650; Orange.
Pfam; PF07527; Hairy_orange; 1.
Pfam; PF00010; HLH; 1.
SMART; SM00353; HLH; 1.
SMART; SM00511; ORANGE; 1.
PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Repressor; Transcription regulation.
FT DNA BIND 45 57 Basic motif.
FT DOMAIN 58 100 Helix-loop-helix motif.
FT DOMAIN 129 175 Orange.
FT DOMAIN 321 373 Ala/Gly-rich.
SQ SEQUENCE 410 AA; 43946 MW; 40A87281B08E233D CRC64;

Query Match 16.7%; Score 81; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 4e-56;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
Db 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102

QY 105 TEQHQKIIALONGERSLKSP 125
Db 103 TEQHQKIIALONGERSLKSP 123

RESULT 6
BHB3_RAT
ID BHB3_RAT STANDARD; PRT; 410 AA.
AC O35779;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (BHLHB3) (Enhancer-of-split and hairy-related protein 1) (SHARP-1).
GN Name=Bhlhb3; Synonyms=Sharp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Cerebellum;
RX MEDLINE=98193761; PubMed=9532582;
RA Rossner M.J., Doerr J., Gass P., Schwab M.H., Nave K.-A.;
RT "SHARPs: mammalian enhancer-of-split- and hairy-related proteins coupled to neuronal stimulation."
RL Mol. Cell. Neurosci. 10:460-475(1997).
CC -!- FUNCTION: May be a transcriptional repressor that represses both basal and activated transcription.
CC -!- SUBUNIT: Homodimerize.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in subregions of the brain, moderately expressed in skeletal muscle, heart. Weakly expressed in lung.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (BHLH) domain.
CC -!- SIMILARITY: Contains 1 orange domain.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a frameshift in position 249.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF009329; AAB63586.1; ALT_FRAME.
RGD; 70900; Bhlhb3.
InterPro; IPR001092; HLH_basic.
InterPro; IPR003650; Orange.
Pfam; PF07527; Hairy_orange; 1.
Pfam; PF00010; HLH; 1.
SMART; SM00353; HLH; 1.
SMART; SM00511; ORANGE; 1.
PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Repressor; Transcription regulation.
FT DNA BIND 45 57 Basic motif.
FT DOMAIN 58 100 Helix-loop-helix motif.
FT DOMAIN 129 175 Orange.
FT DOMAIN 321 344 Ala/Gly-rich.
SQ SEQUENCE 410 AA; 43917 MW; 829705CA3A013127 CRC64;

Query Match 16.7%; Score 81; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 4e-56;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
Db 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102

QY 105 TEQHQKIIALONGERSLKSP 125
Db 103 TEQHQKIIALONGERSLKSP 123

RESULT 7
BAD18970
ID BAD18970 PRELIMINARY; PRT; 410 AA.
AC BAD18970;
DT 20-MAY-2004 (TReMBLrel. 27, Created)
DT 20-MAY-2004 (TReMBLrel. 27, Last sequence update)
DT 20-MAY-2004 (TReMBLrel. 27, Last annotation update)
DE BHLH transcriptional factor Dec2.
GN BHLHB3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11162494;
RA Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K., Yoshida E., Suardita K., Matsuda Y., Kato Y.;
RT "Molecular cloning and characterization of DEC2, a new member of basic helix-loop-helix proteins."
RL Biochem. Biophys. Res. Commun. 280:164-171(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12397359;
RA Honma S., Kawamoto T., Takagi Y., Fujimoto K., Sato F., Noshiro M., Kato Y., Honma K.;
RT "Dec1 and Dec2 are regulators of the mammalian molecular clock."
RL Nature 419:841-844(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Hamaguchi H., Fujimoto K., Kawamoto T., Noshiro M., Maemura K., Takeda N., Nagai R., Furukawa M., Honma S., Honma K., Kurihara H., Kato Y.;
RT "Expression of Dec2 -a basic helix-loop-helix transcription factor- gene is regulated by a molecular clock system."
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB126167; BAD18970.1; --
SQ SEQUENCE 410 AA; 43946 MW; 40A87281B08E233D CRC64;

Query Match 16.7%; Score 81; DB 2; Length 410;
Best Local Similarity 100.0%; Pred. No. 4e-56;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 104
Db 43 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 102

QY 105 TEQQHQKIIALQNGERSLKSP 125
Db 103 TEQQHQKIIALQNGERSLKSP 123

RESULT 8
Q6NY50
ID Q6NY50 PRELIMINARY; PRT; 403 AA.
AC Q6NY50;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein zgc:76913;
GN Name=zgc:76913;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=2338257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC066738; AAH66738.1; -.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy_orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SMO0353; HLH; 1.
DR SMART; SMO0511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
KW Hypothetical protein.
SQ SEQUENCE 403 AA; 44682 MW; 08442DC03FFF298E CRC64;

Query Match 11.0%; Score 53; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 9.3e-34;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 104
Db 43 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 102

QY 105 TEQQHQKIIALQNGERSLKSP 125
Db 103 TEQQHQKIIALQNGERSLKSP 123

RESULT 8
Q6NY50
ID Q6NY50 PRELIMINARY; PRT; 403 AA.
AC Q6NY50;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein zgc:76913;
GN Name=zgc:76913;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=2338257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC066738; AAH66738.1; -.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy_orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SMO0353; HLH; 1.
DR SMART; SMO0511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
KW Hypothetical protein.
SQ SEQUENCE 403 AA; 44682 MW; 08442DC03FFF298E CRC64;

Query Match 11.0%; Score 53; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 9.3e-34;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 97
Db 48 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 100

RESULT 9
AAH66738
ID AAH66738 PRELIMINARY; PRT; 403 AA.
AC AAH66738;
DT 01-JUN-2004 (TREMBlrel. 27, Created)
DT 01-JUN-2004 (TREMBlrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein zgc:76913.
GN ZGC:76913.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC066738; AAH66738.1; -.
KW Hypothetical protein.
SQ SEQUENCE 403 AA; 44682 MW; 08442DC03FFF298E CRC64;

Query Match 11.0%; Score 53; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 9.3e-34;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 97
Db 48 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 100

RESULT 10
Q6GNE9
ID Q6GNE9 PRELIMINARY; PRT; 409 AA.
AC Q6GNE9;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;


```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF010305; AAB64228.1; -.
DR EMBL; Y07836; CAA69169.1; -.
DR EMBL; AF364051; AAK50859.1; -.
DR EMBL; BC010720; AAH10720.1; -.
DR MGD; MGI:1097714; Bhlhb2.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0003677; F:DNA binding; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0016564; F:transcriptional repressor activity; IDA.
DR GO; GO:0045892; P:negative regulation of transcription, DNA-d. . .; IDA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy_orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
FT DNA BIND 53 65 Basic motif.
FT DOMAIN 66 108 Helix-loop-helix motif.
FT DOMAIN 140 184 Orange.
FT CONFLICT 288 288 K -> T (in Ref. 2).
SQ SEQUENCE 411 AA; 45360 MW; B392893CD49292BC CRC64;

Query Match 10.7%; Score 52; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.9e-33;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 TYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKH 97
Db 52 TYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKH 103

RESULT 12
BHB2_RAT STANDARD; PRT; 411 AA.
AC Q35780;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Class B basic helix-loop-helix protein 2 (bHLHB2) (Enhancer-of-split
DE and hairy-related protein 2) (SHARP-2).
GN Name=Bhlhb2; Synonyms=Sharp2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Hippocampus;
RX MEDLINE=98193761; PubMed=9532582;
RA Rossner M.J., Doerr J., Gass P., Schwab M.H., Nave K.-A.;
RT "SHARPs: mammalian enhancer-of-split- and hairy-related proteins
RT coupled to neuronal stimulation.";
RL Mol. Cell. Neurosci. 10:460-475(1997).
CC -!- FUNCTION: May function as a transcriptional factor for neuronal
CC differentiation.
CC -!- SUBUNIT: Able to homodimerize or heterodimerize with E47 (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in heart, brain, spleen, lung,
CC liver, muscle, kidney, uterus and gut.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- SIMILARITY: Contains 1 orange domain.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF009330; AAB63587.1; -.
DR RGD; 68439; Bhlhb2.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy_orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
FT DNA BIND 53 65 Basic motif.
FT DOMAIN 66 108 Helix-loop-helix motif.
FT DOMAIN 140 184 Orange.
SQ SEQUENCE 411 AA; 45529 MW; E56BD468D08824AD CRC64;

Query Match 10.7%; Score 52; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.9e-33;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 TYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKH 97
Db 52 TYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKH 103

RESULT 13
Q76JQ4 PRELIMINARY; PRT; 411 AA.
AC Q76JQ4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Enhancer of split-and hairy-related protein-2.
GN Name=SHARP-2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=14725860;
RA Hirano S., Yamada K., Kawata H., Shou Z., Mizutani T., Shigematsu Y.,
RA Mayumi M., Miyamoto K.;
RT "The rat enhancer of split- and hairy-related protein-2 gene: hepatic
RT expression, genomic structure, and promoter analysis.";
RL Arch. Biochem. Biophys. 422:81-90(2004).
DR EMBL; AB096137; BAD01588.1; -.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy_orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
SQ SEQUENCE 411 AA; 45557 MW; 2086DABB33DB82058 CRC64;

Query Match 10.7%; Score 52; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.9e-33;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 TYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKH 97
Db 52 TYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKH 103

RESULT 14
BAD01588 PRELIMINARY; PRT; 411 AA.
ID BAD01588
AC BAD01588;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2004, 19:56:37 ; Search time 155 Seconds
(without alignments)
1120.161 Million cell updates/sec

Title: US-10-078-650-12
Perfect score: 484
Sequence: 1 MDEGIPHLQERQLLEHRDFI.....PGNPESSAQEDPSQPGKEAP 484

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: _geneseqp1980s:*
2: _geneseqp1990s:*
3: _geneseqp2000s:*
4: _geneseqp2001s:*
5: _geneseqp2002s:*
6: _geneseqp2003as:*
7: _geneseqp2003bs:*
8: _geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	484	100.0	484	4	AAB70693	Aab70693 Human DEC
2	440	90.9	482	4	AAB70692	Aab70692 Human DEC
3	440	90.9	482	5	ABG96308	Abg96308 Human ova
4	440	90.9	482	6	ABO14645	Abol14645 Novel hum
5	440	90.9	482	6	ABR58470	Abr58470 Human bHL
6	440	90.9	482	7	ADC31646	Adc31646 Human nov
7	440	90.9	482	7	ADE54662	Ade54662 Human PRO
8	440	90.9	482	8	ADN05014	Adn05014 Antipsori
9	440	90.9	767	5	ABJ04644	Abj04644 Protein c
10	159	32.9	165	4	AAU16188	Aau16188 Human nov
11	159	32.9	165	6	ABU55257	Abu55257 Human nov
12	108	22.3	203	6	ABO14646	Abol14646 Novel hum
13	107	22.1	150	4	AAU16603	Aau16603 Human nov
14	107	22.1	150	6	ABU55672	Abu55672 Human nov
15	81	16.7	410	4	AAB70694	Aab70694 Mouse DEC
16	81	16.7	410	7	ADE54660	Ade54660 Rat Prote
17	76	15.7	88	4	AAU32415	Aau32415 Novel hum
18	62	12.8	109	5	ADK34819	Adk34819 Novel hum
19	54	11.2	54	5	ABP52802	Abp52802 Medane an
20	52	10.7	379	8	ADP46580	Adp46580 Human col
21	52	10.7	379	8	ADP46579	Adp46579 Human col
22	52	10.7	404	5	ABB57218	Abb57218 Mouse isc
23	52	10.7	411	7	ADE54664	Ade54664 Rat Prote
24	52	10.7	411	7	ADD90602	Add90602 Mouse BHL
25	52	10.7	411	8	ADP74022	Adp74022 Murine DE

26	52	10.7	412	2	AAW99060	Aaw99060 Human bas
27	52	10.7	412	5	ABP65203	Abp65203 Hypoxia-r
28	52	10.7	412	7	ADE54666	Ade54666 Human PRO
29	52	10.7	412	7	ADD90600	Add90600 Human BHL
30	52	10.7	412	8	ADL82901	Adl82901 Human PRO
31	52	10.7	412	8	ADN03704	Adn03704 Antipsori
32	52	10.7	412	8	ADO19143	Adol19143 Human PRO
33	52	10.7	412	8	ADO19145	Adol19145 Human PRO
34	52	10.7	425	8	ADP46583	Adp46583 Human col
35	52	10.7	425	8	ADP46584	Adp46584 Human col
36	49	10.1	207	4	AAU30778	Aau30778 Novel hum
37	47	9.7	54	5	ABP52803	Abp52803 Medane an
38	43	8.9	412	2	AAW08140	Aaw08140 Human cyt
39	43	8.9	412	3	AAy87958	Aay87958 Human CR8
40	43	8.9	412	4	AAB90758	Aab90758 Human she
41	42	8.7	204	7	ADC33242	Adc33242 Human nov
42	28	5.8	432	4	AAU30779	Aau30779 Novel hum
43	16	3.3	18	8	ADH58940	Adh58940 Silk prot
44	16	3.3	19	7	AAO30392	Aao30392 Human ARX
45	16	3.3	20	4	AAB69489	Aab69489 Synthetic

ALIGNMENTS

RESULT 1
AAB70693
ID AAB70693 standard; protein; 484 AA.
XX
AC AAB70693;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human DEC2b protein sequence SEQ ID NO:12.
XX
KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX
OS Homo sapiens.
XX
PN WO200114551-A1.
XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP003991.
XX
PR 19-AUG-1999; 99JP-00233286.
XX
(CHUS) CHUGAI SEIYAKU KK.
XX
PI Fujimoto K, Shin M, Kato Y;
XX
DR WPI; 2001-202935/20.
DR N-PSDB; AAF74776.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation.
XX
PS Claim 2; Page 69-72; 83pp; Japanese.
XX
CC The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in the
CC development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC represents the specifically claimed human DEC2b protein, as given in the
CC present invention
XX
SQ Sequence 484 AA;

Query Match 100.0%; Score 484; DB 4; Length 484;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT and polynucleotides, useful in gene therapy, e.g. for treating or
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
PT stroke or infections.
XX
PS Claim 1; Page 125; 586pp; English.
XX
CC The invention describes a new isolated polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful as therapeutics,
CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This is the amino acid sequence of a novel human NOV protein
XX
SQ Sequence 482 AA;

Query Match 90.9%; Score 440; DB 6; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
DB 43 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102

QY 105 TEQHQKIIALQNGERSLSKSPISQDLDADFHSQFQTCQKEVQLYLSRFESWTPEPRCVQL 164
DB 103 TEQHQKIIALQNGERSLSKSPISQDLDADFHSQFQTCQKEVQLYLSRFESWTPEPRCVQL 162

QY 165 INHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 224
DB 163 INHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 222

QY 225 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 284
DB 223 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 282

QY 285 DSRGGSGGGPGGGGAAAAAALLGPDPAALLLSSLVAFGGGGGAPFPQAAA 344
DB 283 DSRGGSGGGPGGGGAAAAAALLGPDPAALLLSSLVAFGGGGGAPFPQAAA 342

QY 345 AAAPFCLPFCFLSPSAAAAAYVQPFLLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAA 404
DB 343 AAAPFCLPFCFLSPSAAAAAYVQPFLLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAA 402

QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLFPAGPRE 464
DB 403 AAAAAAAPPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLFPAGPRE 462

QY 465 PGNPESSAQEDPSQPGKEAP 484
DB 463 PGNPESSAQEDPSQPGKEAP 482

RESULT 5
ABR58470
ID ABR58470 standard; protein; 482 AA.
XX
AC ABR58470;
XX

DT 07-JUL-2003 (first entry)
XX Human bHLH protein DEC2 from ovarian cancer clone O1668P.
DE
XX Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.
XX
OS Homo sapiens.
XX
PN WO2003029468-A1.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031467.
XX
PR 02-OCT-2001; 2001US-0327135P.
PR 30-MAY-2002; 2002US-0384531P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Mannion J;
XX
DR WPI; 2003-372001/35.
XX
PT New polynucleotide and polypeptide useful for diagnosing and/or treating
PT cancer, particularly ovarian cancer, and as a vaccine.
XX
PS Claim 2; Page 156-157; 169pp; English.
XX
SQ Sequence 482 AA;

Query Match 90.9%; Score 440; DB 6; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
DB 43 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102

QY 105 TEQHQKIIALQNGERSLSKSPISQDLDADFHSQFQTCQKEVQLYLSRFESWTPEPRCVQL 164
DB 103 TEQHQKIIALQNGERSLSKSPISQDLDADFHSQFQTCQKEVQLYLSRFESWTPEPRCVQL 162

QY 165 INHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 224
DB 163 INHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 222

QY 225 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 284
DB 223 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 282

QY 285 DSRGGSGGGPGGGGAAAAAALLGPDPAALLLSSLVAFGGGGGAPFPQAAA 344
DB 283 DSRGGSGGGPGGGGAAAAAALLGPDPAALLLSSLVAFGGGGGAPFPQAAA 342

QY 345 AAAPFCLPFCFLSPSAAAAAYVQPFLLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAA 404
DB 343 AAAPFCLPFCFLSPSAAAAAYVQPFLLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAA 402

QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLFPAGPRE 464
DB 403 AAAAAAAPPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLFPAGPRE 462

QY 465 PGNPESSAQEDPSQPGKEAP 484
DB 463 PGNPESSAQEDPSQPGKEAP 482

RESULT 6
ADC31646
ID ADC31646 standard; protein; 482 AA.
XX.
AC ADC31646;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel polypeptide sequence, SEQ ID NO:1728.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR N-PSDB; ADC30675.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 20; SEQ ID NO 1728; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,

CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 482 AA;
Query Match 90.9%; Score 440; DB 7; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
Db |||||||
QY 43 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102
Db |||||||
QY 105 TEQHQKIIALONGERSLKSPIQSDLDAFHSGFQTCAKEVLQYLSRFESWTPREPRCVQL 164
Db |||||||
QY 103 TEQHQKIIALONGERSLKSPIQSDLDAFHSGFQTCAKEVLQYLSRFESWTPREPRCVQL 162
QY 165 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 224
Db |||||||
QY 163 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 222
QY 225 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDSAPKRMKL 284
Db |||||||
QY 223 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDSAPKRMKL 282
QY 285 DSRGGSGGGPGGGAAGSAAALLGPDPAAGSAAALLRPDAALLSSLVAFGGGGGAPFPQAA 344
Db |||||||
QY 283 DSRGGSGGGPGGGAAGSAAALLGPDPAAGSAAALLRPDAALLSSLVAFGGGGGAPFPQAA 342
QY 345 AAAPFCLPFCFLSPSAAAYVQPFLLDKSGLEKLYPAAAAAPFLLYPGIPAPAAAAAAA 404
Db |||||||
QY 343 AAAPFCLPFCFLSPSAAAYVQPFLLDKSGLEKLYPAAAAAPFLLYPGIPAPAAAAAAA 402
QY 405 AAAAAAAAFPCLSVLSPPPEKAGAGAAATLLPHEVAPLGAPHPQHGHGTHLPFAGPRE 464
Db |||||||
QY 403 AAAAAAAAFPCLSVLSPPPEKAGAGAAATLLPHEVAPLGAPHPQHGHGTHLPFAGPRE 462
QY 465 PGNPESSAQEDPSQPGKEAP 484
Db |||||||
QY 463 PGNPESSAQEDPSQPGKEAP 482
RESULT 7
ADE54662
ID ADE54662 standard; protein; 482 AA.
XX
AC ADE54662;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q9C0J9, SEQ ID NO 467.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 482 AA;

Query Match 90.9%; Score 440; DB 7; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
DB 43 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102
QY 105 TEQHQKIIALQNGERSLSKSPIQSDLDADFHSQFQTCACEVLQYLSRFESWTPREPRCVQL 164
DB 103 TEQHQKIIALQNGERSLSKSPIQSDLDADFHSQFQTCACEVLQYLSRFESWTPREPRCVQL 162
QY 165 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 224
DB 163 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 222
QY 225 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 284
DB 223 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 282
QY 285 DSRGGSGGGPGGGAAGAAALLGPDPAAGAAALLRPDAALLSSLVAFGGGGAPFPQAAA 344
DB 283 DSRGGSGGGPGGGAAGAAALLGPDPAAGAAALLRPDAALLSSLVAFGGGGAPFPQAAA 342
QY 345 AAAPFCLPFCFLSPSAAAAYVQPFLLDKSGLEKLYLPAAAAAPFLLYPGIPAPAAAAAAA 404
DB 343 AAAPFCLPFCFLSPSAAAAYVQPFLLDKSGLEKLYLPAAAAAPFLLYPGIPAPAAAAAAA 402
QY 405 AAAAAAAAFPCLSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLPFAGPRE 464
DB 403 AAAAAAAAFPCLSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGRTHLPFAGPRE 462
QY 465 PGNPESSAQEDPSQPGKEAP 484

DB 463 PGNPESSAQEDPSQPGKEAP 482
RESULT 8
ADN05014
ID ADN05014 standard; protein; 482 AA.
XX AC ADN05014;
XX DT 01-JUL-2004 (first entry)
XX DE Antipsoriatic protein sequence #687.
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX PN WO2004028479-A2.
XX PD 08-APR-2004.
XX PF 25-SEP-2003; 2003WO-US030907.
XX PR 25-SEP-2002; 2002US-0414006P.
XX PA (GETH) GENENTECH INC.
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX DR WPI; 2004-305105/28.
XX DR N-PSDB; ADN05013.
XX PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX PS Claim 9; SEQ ID NO 1408; 3069pp; English.
XX CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX SQ Sequence 482 AA;
Query Match 90.9%; Score 440; DB 8; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
DB 43 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102
QY 105 TEQHQKIIALQNGERSLSKSPIQSDLDADFHSQFQTCACEVLQYLSRFESWTPREPRCVQL 164
DB 103 TEQHQKIIALQNGERSLSKSPIQSDLDADFHSQFQTCACEVLQYLSRFESWTPREPRCVQL 162
QY 165 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 224
DB 163 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 222
QY 225 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 284
DB 223 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGGEDSPAPKRMKL 282
QY 285 DSRGGSGGGPGGGAAGAAALLGPDPAAGAAALLRPDAALLSSLVAFGGGGAPFPQAAA 344
DB 283 DSRGGSGGGPGGGAAGAAALLGPDPAAGAAALLRPDAALLSSLVAFGGGGAPFPQAAA 342
QY 345 AAAPFCLPFCFLSPSAAAAYVQPFLLDKSGLEKLYLPAAAAAPFLLYPGIPAPAAAAAAA 404

Db 343 AAAPFCLPFCFLSPSAAAYVQFFLDKSGLEKYLPAATAAPFLLYPGIPAPAAAAAA 402
QY 405 AAAAAAACPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGHGRTHLPFAGPRE 464
Db 403 AAAAAAACPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGHGRTHLPFAGPRE 462
QY 465 PGNPESSAQEDPSQPGKEAP 484
Db 463 PGNPESSAQEDPSQPGKEAP 482

RESULT 9
ABJ04644
ID ABJ04644 standard; protein; 767 AA.
XX
AC ABJ04644;
XX
DT 11-OCT-2002 (first entry)
XX
DE Protein of NOVX 5 SEQ ID No 10.

XX Cytostatic; antidiabetic; anorectic; metabolic; neuropathic; antilipidemic;
KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
KW tranquiliser; neuroleptic; antidiabetic; antiulcer; antiinflammatory;
KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
KW vesicular transport; cystic fibrosis; gastrointestinal disorder;
KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
XX gene therapy.

OS Unidentified.

XX

XX WO200246409-A2.

PN

XX

PD 13-JUN-2002.

XX

PF 06-DEC-2001; 2001WO-US046586.

XX

PR 06-DEC-2000; 2000US-0251660P.

PR 12-DEC-2000; 2000US-0255029P.

PR 08-JAN-2001; 2001US-0260326P.

PR 24-JAN-2001; 2001US-0263800P.

PR 20-FEB-2001; 2001US-0269942P.

PR 24-APR-2001; 2001US-0286183P.

PR 20-AUG-2001; 2001US-0313627P.

PR 12-SEP-2001; 2001US-0318712P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Guo X, Li L, Patturajan M, Shinkets RA, Casman SJ, Malyankar UM;

PI Tchernev VT, Vernet CAM, Spytek KA, Shenoy SG, Alsobrook JP;

PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA, Boldog FL;

PI Colman SD, Eisen AJ, Liu X, Padigar M, Spaderna SK, Zerhusen BD;

XX

DR WPI; 2002-547774/58.

DR N-PSDB; ABT05457.

XX

XX Novel isolated polypeptide, designated NOVX, useful for treating or

PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and

PT metabolic, neurodegenerative, immune and hematopoietic disorders.

XX

PS Claim 1; Page 44; 421pp; English.

XX

XX The invention relates to an isolated polypeptide, designated NOVX,
CC comprising a sequence fully defined in the specification. The isolated
CC protein, its encoding polynucleotide or an antibody created from the
CC protein is useful in the manufacture of a medicament for treating a

CC syndrome associated with a human disease, preferably a NOVX-associated
CC disorder, or for treating or preventing a NOVX-associated disorder in a
CC subject, preferably human. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are also useful
CC for treating or preventing metabolic disorders, diabetes, obesity,
CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
CC disease, Parkinson's disorder, immune disorders, haematopoietic
CC disorders, and various dyslipidaemias, metabolic disturbances associated
CC with obesity, the metabolic syndrome X, wasting disorders associated with
CC chronic diseases, and cancer. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are useful for
CC treating or preventing neurological disorders such as epilepsy, stroke,
CC mental disorders including mood, anxiety, schizophrenic disorders,
CC disorders of vesicular transport such as cystic fibrosis, diabetes
CC mellitus, goiter, gastrointestinal disorders including ulcerative
CC colitis, other conditions associated with abnormal vesicle trafficking
CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
CC arthritis. A cell comprising the vector of the invention is useful for
CC producing non-human transgenic animals. The polynucleotide of the
CC invention can be used to treat disorders by gene therapy. This sequence
CC represents one of the isolated NOVX proteins of the invention
XX

SQ Sequence 767 AA;

Query Match 90.9%; Score 440; DB 5; Length 767;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLHRIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL 104
Db 328 DTYKLPRLHRIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL 387
QY 105 TEQQHQKIIALQNGERSLKSPIQSDLDAPHSGFQTCAKEVLQYLSRFSWTPREPRCVQL 164
Db 388 TEQQHQKIIALQNGERSLKSPIQSDLDAPHSGFQTCAKEVLQYLSRFSWTPREPRCVQL 447
QY 165 INHLHAVATQFLTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 224
Db 448 INHLHAVATQFLTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 507
QY 225 RTQPSAELAAENDTDTSYGGEGEAERPDREKKGAGASRVTIKQPPGEDSPAPKRMKL 284
Db 508 RTQPSAELAAENDTDTSYGGEGEAERPDREKKGAGASRVTIKQPPGEDSPAPKRMKL 567
QY 285 DSRGGSGGGPGGAAAAAALLGPDPAATAAALLRPDAALLSSLVAFGGGGAPFPQAA 344
Db 568 DSRGGSGGGPGGAAAAAALLGPDPAATAAALLRPDAALLSSLVAFGGGGAPFPQAA 627
QY 345 AAAPFCLPFCFLSPSAAAYVQFFLDKSGLEKYLPAATAAPFLLYPGIPAPAAAAAA 404
Db 628 AAAPFCLPFCFLSPSAAAYVQFFLDKSGLEKYLPAATAAPFLLYPGIPAPAAAAAA 687
QY 405 AAAAAAACPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGHGRTHLPFAGPRE 464
Db 688 AAAAAAACPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAPHPQHGHGRTHLPFAGPRE 747
QY 465 PGNPESSAQEDPSQPGKEAP 484
Db 748 PGNPESSAQEDPSQPGKEAP 767

RESULT 10
AAU16188
ID AAU16188 standard; protein; 165 AA.
XX
AC AAU16188;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human novel secreted protein, Seq ID 1141.
XX
KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;

KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.

XX OS Homo sapiens.

XX WO200155322-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001341.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
DR N-PSDB; AAS26175.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
PS Claim 11; SEQ ID NO 1141; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 32.9%; Score 159; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.4e-132;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGLHLEKAVVLELT LKHLKALTAL 104
Db |||||||||||||||||||||||||||||||||||||||||||||||||||
7 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGLHLEKAVVLELT LKHLKALTAL 66

QY 105 TEQHQKIIALQNGERSLKSPIQSDLDFAHSGFQTCAKEVLQYLSRFESWTPREPRCVQL 164
Db |||||||||||||||||||||||||||||||||||||||||||||||||||
67 TEQHQKIIALQNGERSLKSPIQSDLDFAHSGFQTCAKEVLQYLSRFESWTPREPRCVQL 126

QY 165 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAA 203
Db |||||||||||||||||||||||||||||||||||||||||||||||||||
127 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAA 165

RESULT 11
ABU55257
ID ABU55257 standard; protein; 165 AA.
XX
AC ABU55257;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polypeptide #344.
XX
KW Human; neural disorder; immune system disorder; renal disorder;

KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
XX haemostatic; antiarteriosclerotic.
OS Homo sapiens.
XX
XX US2002132753-A1.
XX
PD 19-SEP-2002.
XX
PF 17-JAN-2001; 2001US-00764864.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-147444/14.

DR N-PSDB; ABX73516.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 11; SEQ ID NO 1141; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX
SQ Sequence 165 AA;

Query Match 32.9%; Score 159; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.4e-132;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
Db 7 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 66
QY 105 TEQHQKIIALQNGERSLKSPISQDLDADFHSQFQTCAKEVLQYLSRFESWTREPRCVQL 164
Db 67 TEQHQKIIALQNGERSLKSPISQDLDADFHSQFQTCAKEVLQYLSRFESWTREPRCVQL 126
QY 165 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAA 203
Db 127 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAA 165

RESULT 12
ABO14646
ID ABO14646 standard; protein; 203 AA.
XX
AC ABO14646;
XX
DT 25-AUG-2003 (first entry)
XX
DE Novel human protein #19.
XX
KW Human; NOV; gene therapy; endocrine related disease; diabetes;
KW metabolism-related disease; obesity; central nervous system disorder;
KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW schizophrenia; depression; autoimmune disease; inflammatory disorder;
KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW stroke; infection.

XX
OS Homo sapiens.
XX
PN WO2003023002-A2.
XX
PD 20-MAR-2003.
XX
PF 09-SEP-2002; 2002WO-US028539.
XX

PR 07-SEP-2001; 2001US-0318120P.
PR 07-SEP-2001; 2001US-0318130P.
PR 10-SEP-2001; 2001US-0318430P.
PR 17-SEP-2001; 2001US-0322636P.
PR 17-SEP-2001; 2001US-0322781P.
PR 17-SEP-2001; 2001US-0322816P.
PR 17-SEP-2001; 2001US-0322817P.
PR 19-SEP-2001; 2001US-0323519P.
PR 20-SEP-2001; 2001US-0323631P.
PR 20-SEP-2001; 2001US-0323636P.
PR 25-SEP-2001; 2001US-0324969P.
PR 25-SEP-2001; 2001US-0325091P.
PR 26-SEP-2001; 2001US-0324990P.
PR 17-APR-2002; 2002US-0373212P.
PR 06-SEP-2002; 2002US-00236177.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
PI Lepley DM, Edinger SR, Burgess CE;
XX
DR WPI; 2003-313242/30.
DR N-PSDB; ACD19339.
XX
PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT and polynucleotides, useful in gene therapy, e.g. for treating or
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
PT stroke or infections.
XX
PS Claim 1; Page 126; 586pp; English.
XX
CC The invention describes a new isolated polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful as therapeutics,
CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This is the amino acid sequence of a novel human NOV protein
XX
SQ Sequence 203 AA;

Query Match 22.3%; Score 108; DB 6; Length 203;
Best Local Similarity 100.0%; Pred. No. 3.1e-87;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
Db 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
QY 105 TEQHQKIIALQNGERSLKSPISQDLDADFHSQFQTCAKEVLQYLSRFE 152
Db 105 TEQHQKIIALQNGERSLKSPISQDLDADFHSQFQTCAKEVLQYLSRFE 152

RESULT 13
AAU16603
ID AAU16603 standard; protein; 150 AA.

XX	AAU16603;			PR	08-SEP-2000;	2000US-0231414P.
AC				PR	08-SEP-2000;	2000US-0232080P.
XX				PR	08-SEP-2000;	2000US-0232081P.
DT	07-NOV-2001 (first entry)			PR	12-SEP-2000;	2000US-0231968P.
XX				PR	14-SEP-2000;	2000US-0232397P.
DE	Human novel secreted protein, Seq ID 1556.			PR	14-SEP-2000;	2000US-0232398P.
XX				PR	14-SEP-2000;	2000US-0232399P.
KW	Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;			PR	14-SEP-2000;	2000US-0232400P.
KW	cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;			PR	14-SEP-2000;	2000US-0232401P.
KW	antibacterial; virucide; fungicide; opthalmological; vulnerary;			PR	14-SEP-2000;	2000US-0233063P.
KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;			PR	14-SEP-2000;	2000US-0233064P.
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;			PR	14-SEP-2000;	2000US-0233065P.
KW	cerebral ischaemia; angiogenesis; nervous system disorder;			PR	21-SEP-2000;	2000US-0234223P.
KW	Alzheimer's disease; infection; ocular disorder; corneal infection;			PR	25-SEP-2000;	2000US-0234997P.
KW	wound healing; epithelial cell proliferation; skin ageing; food additive;			PR	25-SEP-2000;	2000US-0234998P.
KW	preservative; antiproliferative.			PR	26-SEP-2000;	2000US-0235484P.
XX				PR	27-SEP-2000;	2000US-0235834P.
OS	Homo sapiens.			PR	27-SEP-2000;	2000US-0235836P.
XX				PR	29-SEP-2000;	2000US-0236327P.
PN	WO200155322-A2.			PR	29-SEP-2000;	2000US-0236367P.
XX				PR	29-SEP-2000;	2000US-0236368P.
PD	02-AUG-2001.			PR	29-SEP-2000;	2000US-0236369P.
XX				PR	29-SEP-2000;	2000US-0236370P.
PF	17-JAN-2001; 2001WO-US001341.			PR	02-OCT-2000;	2000US-0236802P.
XX				PR	02-OCT-2000;	2000US-0237037P.
PR	31-JAN-2000; 2000US-0179065P.			PR	02-OCT-2000;	2000US-0237038P.
PR	04-FEB-2000; 2000US-0180628P.			PR	02-OCT-2000;	2000US-0237039P.
PR	24-FEB-2000; 2000US-0184664P.			PR	02-OCT-2000;	2000US-0237039P.
PR	02-MAR-2000; 2000US-0186350P.			PR	02-OCT-2000;	2000US-0237040P.
PR	16-MAR-2000; 2000US-0189874P.			PR	13-OCT-2000;	2000US-0239935P.
PR	17-MAR-2000; 2000US-0190076P.			PR	13-OCT-2000;	2000US-0239937P.
PR	18-APR-2000; 2000US-0198123P.			PR	20-OCT-2000;	2000US-0240960P.
PR	19-MAY-2000; 2000US-0205515P.			PR	20-OCT-2000;	2000US-0241221P.
PR	07-JUN-2000; 2000US-0209467P.			PR	20-OCT-2000;	2000US-0241785P.
PR	28-JUN-2000; 2000US-0214886P.			PR	20-OCT-2000;	2000US-0241786P.
PR	30-JUN-2000; 2000US-0215135P.			PR	20-OCT-2000;	2000US-0241787P.
PR	07-JUL-2000; 2000US-0216647P.			PR	20-OCT-2000;	2000US-0241808P.
PR	07-JUL-2000; 2000US-0216880P.			PR	20-OCT-2000;	2000US-0241809P.
PR	11-JUL-2000; 2000US-0217487P.			PR	20-OCT-2000;	2000US-0241826P.
PR	11-JUL-2000; 2000US-0217496P.			PR	01-NOV-2000;	2000US-0244617P.
PR	14-JUL-2000; 2000US-0218290P.			PR	08-NOV-2000;	2000US-0246474P.
PR	26-JUL-2000; 2000US-0220963P.			PR	08-NOV-2000;	2000US-0246475P.
PR	26-JUL-2000; 2000US-0220964P.			PR	08-NOV-2000;	2000US-0246476P.
PR	14-AUG-2000; 2000US-0224518P.			PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000; 2000US-0224519P.			PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000; 2000US-0225213P.			PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000; 2000US-0225214P.			PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000; 2000US-0225266P.			PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000; 2000US-0225267P.			PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000; 2000US-0225268P.			PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000; 2000US-0225270P.			PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000; 2000US-0225447P.			PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000; 2000US-0225757P.			PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000; 2000US-0225758P.			PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000; 2000US-0225759P.			PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000; 2000US-0226279P.			PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000; 2000US-0226681P.			PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000; 2000US-0226868P.			PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000; 2000US-0227182P.			PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000; 2000US-0227009P.			PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000; 2000US-0228924P.			PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000; 2000US-0229287P.			PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000; 2000US-0229343P.			PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000; 2000US-0229344P.			PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000; 2000US-0229345P.			PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000; 2000US-0229509P.			PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000; 2000US-0229513P.			PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000; 2000US-0230437P.			PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000; 2000US-0230438P.			PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000; 2000US-0231242P.			PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000; 2000US-0231243P.			PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000; 2000US-0231244P.			PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000; 2000US-0231413P.			PR	17-NOV-2000;	2000US-0249297P.

PA (BARA/) BARASH S C.
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-147444/14.
DR N-PSDB; ABX73931.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 11; SEQ ID NO 1556; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX
SQ Sequence 150 AA;

Query Match 22.1%; Score 107; DB 6; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.8e-86;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 LGHLEKAVVLELT LKHLKALTALTEQOQKIIALQNGERSLKSP IQSDLD AFHSGFQCA 141
Db 44 LGHLEKAVVLELT LKHLKALTALTEQOQKIIALQNGERSLKSP IQSDLD AFHSGFQCA 103

QY 142 KEVLQYLSRFESWTPREPRCVQLINHLHAVATQFLPTPQLLTQQVPL 188
Db 104 KEVLQYLSRFESWTPREPRCVQLINHLHAVATQFLPTPQLLTQQVPL 150

RESULT 15
AAB70694
ID AAB70694 standard; protein; 410 AA.
XX
AC AAB70694;
XX
DT 17-MAY-2001 (first entry)
XX
DE Mouse DEC2a protein sequence SEQ ID NO:14.
XX
KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX
OS Mus musculus.
XX
PN WO200114551-A1.
XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP003991.
XX
PR 19-AUG-1999; 99JP-00233286.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Fujimoto K, Shin M, Kato Y;
XX
DR WPI; 2001-202935/20.

DR N-PSDB; AAP74777.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation.
XX
PS Claim 2; Page 77-80; 83pp; Japanese.
XX
CC The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in the
CC development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC represents the specifically claimed mouse DEC2a protein, as given in the
CC present invention
XX
SQ Sequence 410 AA;

Query Match 16.7%; Score 81; DB 4; Length 410;
Best Local Similarity 100.0%; Pred. No. 5.3e-63;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELT LKHLKALTAL 104
Db 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELT LKHLKALTAL 102

QY 105 TEQOQKIIALQNGERSLKSP 125
Db 103 TEQOQKIIALQNGERSLKSP 123

Search completed: December 15, 2004, 20:10:58
Job time : 164 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2004, 19:36:35 ; Search time 41 Seconds
(without alignments)
1135.827 Million cell updates/sec

Title: US-10-078-650-12
Perfect score: 2517
Sequence: 1 MDEGIPHLQERQLLEHRDFI.....PGNPSSAQEDPSQPGKEAP 484

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2498	99.2	482	2 JC7583	basic helix-loop-h
2	1634	64.9	410	2 JC7584	basic helix-loop-h
3	669	26.6	412	2 JC5547	basic helix-loop-h
4	209.5	8.3	460	2 T33110	hypothetical prote
5	188	7.5	218	2 S29712	segmentation prote
6	187	7.4	280	2 A53027	transcription fact
7	183.5	7.3	281	2 S36748	transcription fact
8	182.5	7.3	435	2 A46231	helix-loop-helix p
9	182	7.2	581	2 T22341	hypothetical prote
10	181	7.2	465	2 G02738	FRAC-4 - human
11	180	7.2	418	2 T15142	hypothetical prote
12	179.5	7.1	676	1 EDBE23	immediate-early pr
13	177	7.0	282	2 A53336	transcription fact
14	176	7.0	846	2 S52418	GTP-binding regula
15	175.5	7.0	627	2 A44112	spidroin 2, dragli
16	172.5	6.9	416	1 SKXLAG	dermal gland prote
17	171.5	6.8	420	2 A49642	transcription fact
18	171.5	6.8	676	1 EDBE22	immediate-early pr
19	171.5	6.8	801	2 T29018	hypothetical prote
20	171.5	6.8	3534	2 T42567	tegument protein 2
21	170.5	6.8	825	1 EDBEXD	immediate-early pr
22	170	6.8	901	2 A49227	sialidase - Actino
23	169	6.7	702	2 G01840	T-box protein 2 -
24	169	6.7	862	2 T46289	hypothetical prote
25	168.5	6.7	416	2 S12541	evx1 protein - mur
26	167.5	6.7	903	2 T00705	N-chimerin homolog
27	167	6.6	780	2 T00366	hypothetical prote
28	167	6.6	2639	2 T31328	fibroin - Chinese
29	164.5	6.5	1334	2 T50568	probable multi-dom

30	163.5	6.5	839	2 F75518	hypothetical prote
31	163	6.5	3938	2 T42761	Bassoon protein -
32	162.5	6.5	1077	2 A44067	serine-rich protei
33	162.5	6.5	1585	2 T31611	hypothetical prote
34	162	6.4	337	2 S06956	segmentation prote
35	161.5	6.4	443	1 I38239	transcription fact
36	160.5	6.4	573	1 S20710	hypothetical prote
37	159.5	6.3	205	2 B46177	enhancer of split
38	159.5	6.3	701	1 S46458	transcription fact
39	159	6.3	660	1 QQBE3	BHLF1 protein - hu
40	157.5	6.3	901	2 A44825	phosphoprotein, sy
41	157	6.2	195	2 C46177	enhancer of split
42	157	6.2	378	2 A44443	basic helix-loop-h
43	157	6.2	860	1 EAMS	elastin precursor
44	157	6.2	1119	2 TS0995	related to cytoske
45	157	6.2	1952	2 T48814	hypothetical prote

ALIGNMENTS

RESULT 1

JC7583

basic helix-loop-helix protein, DEC2 - human

C/Species: Homo sapiens (man)

C/Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004

C/Accession: JC7583

R/Fujimoto, K.; Shen, M.; Noshiro, M.; Matsubara, K.; Shingu, S.; Honda, K.; Yoshida, E. Biochem. Biophys. Res. Commun. 280, 164-171, 2001

A/Title: Molecular cloning and characterization of DEC2, a new member of basic helix-loop-loo

A/Reference number: JC7583; MUID:21092582; PMID:11162494

A/Accession: JC7583

A/Molecule type: mRNA

A/Residues: 1-482 <FUJ>

A/Cross-references: UNIPROT:Q9C0J9; DDBJ:AB044088

C/Comment: This protein, a novel member of the DEC subfamily of basic helix-loop-helix p

C/Genetics:

A/Gene: dec2

A/Map position: 12p11.23-p12.1

C/Keywords: transcription factor

F;1-173/Region: highly conserved #status predicted

F;130-173/Domain: Orange #status predicted <ORA>

F;286-411/Region: alanine and glycine-rich #status predicted

Query Match	99.2%;	Score	2498;	DB	2;	Length	482;
Best Local Similarity	99.6%;	Pred. No.	4.8e-140;				
Matches	482;	Conservative	0;	Mismatches	0;	Indels	2;
						Gaps	1;

QY	1	MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDDTKVSDTYKLPRLIEKKRRD	60
Db	1	MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDDTK--DTYKLPRLIEKKRRD	58

QY	61	RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTALTEQQHQKIALLQNGER	120
Db	59	RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTALTEQQHQKIALLQNGER	118

QY	121	SLKSPIQSDLDAPHSGFQTCAKEVLYLSRFESWTPREPRCVQLINHLHAVATQFLPTPQ	180
Db	119	SLKSPIQSDLDAPHSGFQTCAKEVLYLSRFESWTPREPRCVQLINHLHAVATQFLPTPQ	178

QY	181	LLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT	240
Db	179	LLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT	238

QY	241	DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDSAPAKRMKLDSSRGSGGGGGGAA	300
Db	239	DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDSAPAKRMKLDSSRGSGGGGGGAA	298

QY	301	AAAAALLGPDAAAAALLRPDAALLSSILVAFGGGGGAPFPQPAATAAPFCFLSPSA	360
Db	299	AAAAALLGPDAAAAALLRPDAALLSSILVAFGGGGGAPFPQPAATAAPFCFLSPSA	358

QY	361	AAAYVQPFLLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAAFAFPCLSSV	420
----	-----	---	-----

Db 359 AAAYQPFLEKSGLEKYLPAAPAAAPFLLYPGIPAPAAAAAFAAPCLSSV 418
QY 421 LSPPEKAGAAATLLPHEVAPLGAPHPQHPGHRTHLPFAGPREPGNPESQAQEDPSQPG 480
Db 419 LSPPEKAGAAATLLPHEVAPLGAPHPQHPGHRTHLPFAGPREPGNPESQAQEDPSQPG 478
QY 481 KEAP 484
Db 479 KEAP 482

RESULT 2
JC7584
basic helix-loop-helix protein, DEC2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C:Accession: JC7584
R:Fujimoto, K.; Shen, M.; Noshiro, M.; Matsubara, K.; Shingu, S.; Honda, K.; Yoshida, E.
Biochem. Biophys. Res. Commun. 280, 164-171, 2001
A:Title: Molecular cloning and characterization of DEC2, a new member of basic helix-loop-helix protein family of the DEC subfamily of basic helix-loop-helix proteins
A:Reference number: JC7583; MUID:21092582; PMID:11162494
A:Accession: JC7584
A:Molecule type: mRNA
A:Residues: 1-410 <FUJ>
A:Cross-references: UNIPROT:Q99PV5; DDBJ:AB044090
C:Comment: This protein, a novel member of the DEC subfamily of basic helix-loop-helix proteins
C:Genetics:
A:Gene: dec2
A:Map position: 6 G2-G3
C:Keywords: transcription factor
F:1-173/Region: highly conserved #status predicted
F:130-173/Domain: Orange #status predicted <ORA>
F:277-344/Region: alanine and glycine-rich #status predicted

Query Match 64.9%; Score 1634; DB 2; Length 410;
Best Local Similarity 72.0%; Pred. No. 3.2e-89;
Matches 350; Conservative 21; Mismatches 37; Indels 78; Gaps 16;

QY 1 MDEGPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDITKVSPTYKLPRLIEKKRRD 60
Db 1 MDEGPHLQERQLLEHRDFIGLDYSSLYMCKPKRSLKDDTK--DTYKLPRLIEKKRRD 58
QY 61 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTALTEQHQHKLIALQNGER 120
Db 59 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTALTEQHQHKLIALQNGER 118
QY 121 SLKSPISQDLDAFHSGFQTCACEVLYLSRFESWTREPRCQVQLINHLHVAATQFLPTPQ 180
Db 119 SLKSPVQADLDAFHSGFQTCACEVLYLSRFESWTREPRCAQLVSHLHVAAT--Q 172
QY 181 LLTQVPLSKGTG-APSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTD 239
Db 173 LLTPQVPSGRGSRAPCSAGAAASGPER-----VACVPIQRTQPGTE--PEHDTD 223
QY 240 TDSYGGEAEAPDREKKGAGASRVTIKQEPPEGDSAPKRMKLDNRGGSGGGPGGGA 299
Db 224 TDSYGGEAE-----QGRAAVKQEPPEGDSAPKRPKLEARG----- 260
QY 300 AAAAAALLGPDPAAPAAALLRPDAALLSSLVAFGGGGAPFPQAAAAAPFLPFCFLSPS 359
Db 261 -----ALLGPEP-----ALLGSLVAL--GGGAPFAQP--AAAPFLPFLYLLSPS 300
QY 360 AAAAYQPFLEKSGLEKYLPAAPAAAPFLLYPGIPAPAAAAAFAAPCLSS 419
Db 301 -AAAYQVPLDKSGLDKYLPAAPAAAPFLLYPGIP-----AAAAAFAAPCLSS 350
QY 420 VLSPPPEKAGA-AAATLLPHEVAPLGAPHPQHPGHRTHLPFAGPREPGNPESQAQEDPSQ 478
Db 351 VLSPPPEKAGATAGAPFLAHEVAPPGPLRPQHAHSRTHLPRA-----VNPESQ-QEDATQ 404
QY 479 PGKEAP 484

Db 405 PAKDAP 410

RESULT 3
JC5547
basic helix-loop-helix factor DEC1 - human
C:Species: Homo sapiens (man)
C:Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 09-Jul-2004
C:Accession: JC5547
R:Shen, M.; Kawamoto, T.; Yan, W.; Nakamasu, K.; Tamagami, M.; Koyano, Y.; Noshiro, M.; R.
Biochem. Biophys. Res. Commun. 236, 294-298, 1997
A:Title: Molecular characterization of the novel basic helix-loop-helix protein DEC1
A:Reference number: JC5547; MUID:97382424; PMID:9240428
A:Accession: JC5547
A:Molecule type: mRNA
A:Residues: 1-412 <SHE>
A:Cross-references: UNIPROT:O14503; DDBJ:AB004066; NID:g2308996; PIDN:BAA21720.1; PID:g2308996
C:Comment: This protein is involved in the control of cell differentiation in several tissues
F:51-108/Region: helix-loop-helix #status predicted

Query Match 26.6%; Score 669; DB 2; Length 412;
Best Local Similarity 40.7%; Pred. No. 2.2e-32;
Matches 191; Conservative 55; Mismatches 93; Indels 130; Gaps 25;

QY 14 LEHRDFIGLDYSSLY-MCKPKRSMKR-DDTKVSDTYKLPRLIEKKRRDRINECIAQLKD 71
Db 20 LEHGDLPGMYPAHMYQVYKSRGKRSSEDSK--ETKLPRLIEKKRRDRINECIAQLKD 77
QY 72 LLPEHLKLTTLGHLEKAVVLELTTLKHLKALTALTEQHQHKLIALQN----GERSLKSPIQ 127
Db 78 LLPEHLKLTTLGHLEKAVVLELTTLKHLKALTALTEQHQHKLIALQSGLOAGELSGRN-VE 136
QY 128 SLDLAFHSGFQTCACEVLYLSRFESWTREPRCQVQLINHLHVAATQFL-----P 177
Db 137 TGQEMFCSGFQTCAREVLQYLAKHEN--TRDLKSSQLVTHLHRVSELQGTSTRKPSDP 194
QY 178 TPQLLT-QQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRT--QPSAELAA 234
Db 195 APKVMDFEKEKPSAPKSGSE-----PGKNCVPIQRTFAHSSGEQSG 236
QY 235 ENDTDTDSYGGEAE---ARPDREKKGAGASRVTI-----IKQEPPEGDSAPK-RMK 283
Db 237 -SDTDTDSYGGESEKGLRSEQPCFKSDHGRFTMGERIGAIKQE--SEEPPTKXNRMQ 293
QY 284 L-DSRGGSGGGPGGGAAGAAALLGPDPAAPAAALLRPDAALLSSLVAFGGGGAPFPQP 342
Db 294 LSDDEGHFTS-----SDLLS-----SPFLGP 314
QY 343 AAAAAAPFLPFCFLSPSAAAAAYQVPLDKSGLEKYLPAAPAAAPFLLYPGIPAPAAAA 402
Db 315 HPHQPPFCLPF-YLIPPSATAYL-PM-----LEKQWYPTSV-----PVLYPGLNASAAA-- 361
QY 403 AAAAAAFAAPCLSSVLSPPEKAGAAAAATLLPHEVAPLGAPHPQHP 451
Db 362 -----LSSFMNPDK-----ISAPLLMPQR---LPSPLPAHP 389

RESULT 4
T33110
hypothetical protein C18H7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33110
R:Tin-Wollam, A.; Fronick, W.
Submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid C18H7.
A:Reference number: Z21284
A:Accession: T33110
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-460 <TIN>
A:Cross-references: UNIPROT:Q9GZF7; EMBL:AF067607; PIDN:AAC17641.1; GSPDB:GN00022; CESP:C
A:Experimental source: strain Bristol N2; clone C18H7

C;Genetics:
A;Gene: CBSP:C18H7.3
A;Map position: 4
A;Introns: 84/1
C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match
Best Local Similarity 8.3%; Score 209.5; DB 2; Length 460;
Matches 99; Conservative 15; Mismatches 119; Indels 81; Gaps 14;

QY 193 GAPSAAGS---AAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDTDSGYGGEAE 249
DB 186 GKPGAAGSDGEACAPGTTKTNPAGEP-----GPAGPPGPPGPNGEAGAGSGPGPEGP 237

QY 250 ARPDREKKG--AGASRVTIKQEPPEGEDSPAPKRMKLDNRGGSGGGPGGGAA----- 300
DB 238 AGPPGPNKDGGEAGA-----PGKDGDA-----GADGAPGTDAAAYCPCPPR 277

QY 301 -AAAAALLGPDPA---AALLRPDAALLSSLVAFGGGGGAPFPQPA---AAPPCLPFCEFL 356
DB 278 SAALGAGGGAEPAGAAPEAAAPAAPEAAPEAAPEAAPEAAPEAAPEAAPEAAPEAAPEAA 331

QY 357 SPSAAAAYVQFLDKSGLEKYLYPAA--AAAPFLLYPGIPAPAPAAAAA-----AAAF 414
DB 332 APEAAPAAEAPAAEAGAGGAEPAAGADPDAAPAAAPAAEAPAAEAPAAEAGGGAEPAGAAA- 390

QY 415 PCLSSVLSPPEKAGAAAAATLLPHEVAPL--GAPHPQHGHGTHLPFAGPREPGN--PES 470
DB 391 -----PDAAAAAPEAAAAPEAAPEAAPEAAPEAAPEAAPEAAPEAAPEAAPEAAPEAA 426

QY 471 SAQEDPSQPGKEAP 484
DB 427 AAAAPEGAGGGEAP 440

RESULT 5
S29712
segmentation protein hairy - red flour beetle
C;Species: Tribolium castaneum (red flour beetle)
C;Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 09-Jul-2004
C;Accession: S29712
R;Sommer, R.J.; Tautz, D.
Nature 361, 448-450, 1993
A;Title: Involvement of an orthologue of the Drosophila pair-rule gene hairy in segment
A;Reference number: S29712; MUID:93156810; PMID:8429884
A;Accession: S29712
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-218 <SOM>
A;Cross-references: UNIPROT:Q8T306

Query Match
Best Local Similarity 7.5%; Score 188; DB 2; Length 218;
Matches 72; Conservative 28; Mismatches 91; Indels 48; Gaps 10;

QY 53 LIEKRRDRINECIAQLKDLLPEHLKLTTLGH--LEKAVVLELTLKHLKALTALTEQHQ 110
DB 5 IMEKRRRRARINNSLNLKTLILDAMKKDPAHSHKLEKADILENTVKHLQNL-----QRQ 58

QY 111 KIIALQNGERSLKSPIQSDLDFAHSGFQTCAKEVLQYLSRFESWTPREPRCVQLINHLHA 170
DB 59 QAAMWQPTDPSVVS-----KFRAGFSECASEV---GRFPGLDPVVKR--RLQLHLAS 105

QY 171 VATQFLPTPQ-----LLTQQVPLSKGTGAP-----SAAGSAAAPCLERAGQKLEPLAY 218
DB 106 CLNQGQKEPQVQVIVPEVAPNNIILNGTGVQLVPIRLANGDIA---LVLPTQGASPLPL 162

QY 219 CVPVIQRTQPSAELAAENDTDTDSGYGGEAEARPDREKKGKAGASRVTIKQEPPEGDSP 277
DB 163 LVPIPQRTASTASSA-----SNYSPSQSPESVR-----PLSLVVRRRREPTTEKP 209

RESULT 6

A53027
transcription factor HES-1 - human
N;Alternate names: hairy protein homolog
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A53027
R;Feder, J.N.; Li, L.; Jan, L.Y.; Jan, Y.N.
Genomics 20, 56-61, 1994
A;Title: Genomic cloning and chromosomal localization of HRY, the human homolog to the D.
A;Reference number: A53027; MUID:94292187; PMID:8020957
A;Accession: A53027
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <FED>
A;Cross-references: UNIPROT:Q14469; GB:L19314; NID:9436999; PIDN:AAA65220.1; PID:G780128
C;Genetics:
A;Gene: GDB:HRY
A;Cross-references: GDB:217077; OMIM:139605
A;Map position: 3q28-3q29
C;Keywords: immediate-early protein

Query Match
Best Local Similarity 7.4%; Score 187; DB 2; Length 280;
Matches 86; Conservative 25; Mismatches 88; Indels 114; Gaps 14;

QY 31 KPKRSMKRDDTKVSDTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGH--LEKA 88
DB 27 KPKTA-----SEHRKSSKPIMEKRRRARINESLSQLKTLILDALKDSSRHSKLEKA 78

QY 89 VLELTLKHLKALTALTEQHQKIIALQNGERSLKSPIQSDLDFAHSGFQTCAKEVLQYL 148
DB 79 DILEMTVKHLRNL-----QRAQMTAAL-----STDPSVLGKYRAGFSECNEVTRFL 125

QY 149 SRFESWTPREPRCVQLINHLHAVATQF-----LPTPQLLTQQVPLSKGTGAPSAAGSA 201
DB 126 STCEG-VNTEVR-TRLGLHLANCMQTINAMTYPGQHPALQAPPPP-PGPGGPGQHAPFA 182

QY 202 AAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAG 261
DB 183 PPP-----PL---VPI-----PDPAAA 190

QY 262 ASRVTIKQEPPEGEDSPAPKRMKLDNRGGSGGGPG--GGAAAAAALLG-----PDPAAA 314
DB 191 -----PGGAAPP-----GGAPCKLGSQAGEAAKVFQGVVVPAPDQG 228

QY 315 AALLRPDAALLSS 327
DB 229 FAFLIPNGAFAPHS 241

RESULT 7
S36748
transcription factor HES-1 - rat
N;Alternate names: hairy-like gene RHL; helix-loop-helix protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S36748; A48106
R;Sasai, Y.; Kageyama, R.; Tagawa, Y.; Shigemoto, R.; Nakanishi, S.
Genes Dev. 6, 2620-2634, 1992
A;Title: Two mammalian helix-loop-helix factors structurally related to Drosophila hairy
A;Reference number: S36748; MUID:94040724; PMID:1340473
A;Accession: S36748
A;Molecule type: mRNA
A;Residues: 1-281 <SAS>
A;Cross-references: UNIPROT:Q04666; EMBL:D13417; NID:9220760; PIDN:BAA02682.1; PID:d1003
R;Feder, J.N.; Jan, L.Y.; Jan, Y.N.
Mol. Cell. Biol. 13, 105-113, 1993
A;Title: A rat gene with sequence homology to the Drosophila gene hairy is rapidly induc
A;Reference number: A48106; MUID:93109293; PMID:8417318
A;Accession: A48106
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-85,'N',87-92,'W',94-281 <FED>


```

Query Match      7.0%; Score 177; DB 2; Length 282;
Best Local Similarity 26.7%; Pred. No. 0.0014;
Matches 83; Conservative 25; Mismatches 95; Indels 108; Gaps 12;

QY   31 KPKRSMKRDDTKVSDTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTLGLH--LEKA 88
    ||| :          :|||:|||:|||:|||:|||:|||:|||:|||:|||
Db   27 KPKTA-----SEHRKSSKPIMEKRRRARINESLSQLKTLLILDALKDSSRHSKLEKA 78
    ||| :          :|||:|||:|||:|||:|||:|||:|||:|||:|||

QY   89 VVLELTCLKHALTALTEQQHQKIIALONGERSLKSPIQSDDLDAFHSGFQTCAKEVLQYL 148
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db   79 DILEMTVKHLRNL-----QRAQMTAAL-----STDPSVLGKYRAGFSECNNEVTREL 125
    :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY   149 SRFESWTPPREPRVCVOLINHHLHAVATQF-----LPTPQLLTQQVPLSKGTCAPSAAGSA 201
    |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db   126 STCEG-VNTEVR-TRLGLHLANCMTQINAMTYPGQAHPALQAPPPPPSGEAGPQHAPFA 183
    |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY   202 AAPCLERAGQKLEPLAYCVPVIQTQPSEAELAENDTDTSYGYGGEAEARPDREKKGKAG 261
    |         ||     |||
Db   184 PPP-----PPL---VPI-----192
    |         ||     |||

QY   262 ASRVTIKQEPGEDSPAPKRMKLDNRGGGGGGGGAATAAAALLG-----PDPAASAAA 316
    |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db   193 -----PGGAAPPP-----GSAPCKLGSQLAGEAAKVFGGVFPVAPADGQFA 232
    |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY   317 LLRPDAALLSS 327
    |||:|||:|||
Db   233 FLIPNGAFSAHS 243

RESULT 14
S52418
GTP-binding regulatory protein Gs alpha-XL chain - rat
N;Alternate names: G protein XL-alpha-s
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
C;Accession: S52418
R;Kehlenbach, R.H.; Matthey, J.; Huttner, W.B.
Nature 372, 804-809, 1994
A;Title: XL-alpha-s is a new type of G protein.
A;Reference number: S52418; PMID:95089824; PMID:7997272
A;Accession: S52418
A;Molecule type: mRNA
A;Residues: 1-846 <KEH>
A;Cross-references: EMBL:X84047; NID:g6422267; PIDN:CAA58866.1; PID:g6422268
R;Kehlenbach, R.H.; Matthey, J.; Huttner, W.B.
Nature 375, 253, 1995
A;Title: Correction: XLaiphas is a new type of G protein.
A;Reference number: S58911
A;Contents: annotation; assignment of start_codon
A;Note: experimental data from this paper suggest that the translation is initiated at p
C;Keywords: GTP binding; nucleotide binding; P-loop; signal transduction
F;132-846/Product: GTP-binding regulatory protein Gs alpha-XL chain #status experimental
F;499-506/Region: nucleotide-binding motif A (P-loop)
F;744-747/Region: GTP-binding NKXD motif

```

[illegible]

```

Query Match      7.0%; Score 175.5; DB 2; Length 627;
Best Local Similarity 25.9%; Pred. No. 0.0041;
Matches 109; Conservative 20; Mismatches 135; Indels 157; Gaps 18;

QY 184 QVPLSKGTG--APSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDTD 241
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 93  QQGGPGYGQGGQGGPSGPGSAAA-----ASAAASAESGQQGP 128

QY 242 SGYG-----GEAEARPDREKKGAG-ASRVTIKQEPGGEDSP---APKRMK 283
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 129 GGYGPGQQGGPGYGPGQQGGPGYGPGQQGGPSGSAASAAAAASGPGQQGGPGYGPGQQG 188

QY 284 LDSRGGGSGG--GPGGGAASAAAA-----LLGPDPAAA 314
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 189 PGGYGPQQGGPSGPGSAAAAASGPGQQGGPGYGPGQQGGPGYGPGQQGSLGPGSAAA 248

QY 315 AALL-----RPDAALLSSLVAFGGGGGAPFPQPAASAAAPFCLPFCFL 356
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 249 AAAAGPGQQGGPGYGPGQQGPGSPGSAASAAAAAAGPGYGPGQQGGPGYGPGQQGPGSGA 308

QY 357 SPSAAAAAYQFLLDKSGLEKY-----LYPAAAAAP---FP 388
      :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 309 GSAAAAAAAAGP--GQQGLGGYGPQQGGPGYGPGQQGGPGYGPGSASAAAAAAGPGQQGP 366

QY 389 LLY-PGIPAPAAAAASAAAAASAAAA-----FPCLSSVLSPPEKAGAAAA 433
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 367 GGYGPGQQGGPSGPGSASAAAAAAGPGYGPGQQGGPGYGAPGOOGPGSPGSAASAAAAA 426

```

QY 434 TLLPHEVAPLG-APHPOHPHGRTHLP-FAGPREPGNPES-----AQEDPSQP 479
Db 427 A-----AGPGGYGPGQQGPGG--YAPCQQGPGSGPSAAAAAGPGGYGPAQQGPGSP 479
QY 480 G 480
Db 480 G 480

Search completed: December 15, 2004, 19:56:33
Job time : 44 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2004, 19:15:04 ; Search time 154 Seconds
(without alignments)
1127.435 Million cell updates/sec

Title: US-10-078-650-12
Perfect score: 2517
Sequence: 1 MDEGIPHLQERQLLEHRDFI.....PGNPESSAQEDPSQPGKEAP 484

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2517	100.0	484	4 AAB70693	Aab70693 Human DEC
2	2498	99.2	482	4 AAB70692	Aab70692 Human DEC
3	2498	99.2	482	5 ABG96308	Abg96308 Human ova
4	2498	99.2	482	6 ABO14645	Abol14645 Novel hum
5	2498	99.2	482	6 ABR58470	Abr58470 Human bHL
6	2498	99.2	482	7 ADC31646	Adc31646 Human nov
7	2498	99.2	482	7 ADE54662	Ades4662 Human Pro
8	2498	99.2	482	8 ADN05014	Adn05014 Antipsori
9	2498	99.2	767	5 ABJ04644	Abj04644 Protein o
10	1634	64.9	410	4 AAB70694	Aab70694 Mouse DEC
11	1624	64.5	410	7 ADE54660	Ades4660 Rat Prote
12	905	36.0	203	6 ABO14646	Abol14646 Novel hum
13	835	33.2	165	4 AAU16188	Aau16188 Human nov
14	835	33.2	165	6 ABU55257	Abu55257 Human nov
15	750	29.8	150	4 AAU16603	Aau16603 Human nov
16	750	29.8	150	6 ABU55672	Abu55672 Human nov
17	685.5	27.2	404	5 ABB57218	Abb57218 Mouse isc
18	685.5	27.2	411	7 ADD90602	Add90602 Mouse BHL
19	685.5	27.2	411	8 ADP74022	Adp74022 Murine DE
20	682.5	27.1	411	7 ADE54664	Ades4664 Rat Prote
21	669	26.6	412	2 AAW08140	Aaw08140 Human cyt
22	669	26.6	412	2 AAW99060	Aaw99060 Human bas
23	669	26.6	412	3 AAY87958	Aay87958 Human CR8
24	669	26.6	412	4 AAB90758	Aab90758 Human she
25	669	26.6	412	5 ABP65203	Abp65203 Hypoxia-r

26	669	26.6	412	7 ADE54666	Ades4666 Human Pro
27	669	26.6	412	7 ADD90600	Add90600 Human BHL
28	669	26.6	412	8 ADL82901	Adl82901 Human PRO
29	669	26.6	412	8 ADN03704	Adn03704 Antipsori
30	669	26.6	412	8 ADO19143	Ado19143 Human PRO
31	669	26.6	412	8 ADO19145	Ado19145 Human PRO
32	669	26.6	425	8 ADP46583	Adp46583 Human col
33	669	26.6	425	8 ADP46584	Adp46584 Human col
34	638	25.3	379	8 ADP46580	Adp46580 Human col
35	638	25.3	379	8 ADP46579	Adp46579 Human col
36	606.5	24.1	432	4 AAU30779	Aau30779 Novel hum
37	458.5	18.2	88	4 AAU32415	Aau32415 Novel hum
38	417.5	16.6	109	5 ADK34819	Adk34819 Novel hum
39	404	16.1	204	7 ADC33242	Adc33242 Human nov
40	333.5	13.2	207	4 AAU30778	Aau30778 Novel hum
41	269	10.7	54	5 ABP52802	Abp52802 Medane an
42	260	10.3	54	5 ABP52803	Abp52803 Medane an
43	220.5	8.8	339	4 AAB72376	Aab72376 Murine ca
44	220	8.7	390	3 AAB42083	Aab42083 Human ORF
45	207	8.2	337	4 AAB47269	Aab47269 Human gri

ALIGNMENTS

RESULT 1
AAB70693
ID AAB70693 standard; protein; 484 AA.
XX
AC AAB70693;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human DEC2b protein sequence SEQ ID NO:12.
XX
KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX
OS Homo sapiens.
XX
PN WO200114551-A1.
XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP003991.
XX
PR 19-AUG-1999; 99JP-00233286.
XX
(CHUS) CHUGAI SEIYAKU KK.
Fujimoto K, Shin M, Kato Y;
WPI; 2001-202935/20.
DR N-PSDB; AAF74776.
XX
DEC2 is a basic helix loop helix protein of the DEC family for use in
development of drugs for treatment of disorders of cell differentiation
and proliferation.
XX
Claim 2; Page 69-72; 83pp; Japanese.
XX
The present invention describes a basic helix loop helix (bHLH) type
transcription factor designated DEC2. DEC2 can be used as a tool in the
development of drugs for the treatment and prevention of disorders
involving cell differentiation and proliferation. The present sequence
represents the specifically claimed human DEC2b protein, as given in the
present invention
XX
Sequence 484 AA;
XX
Query Match 100.0%; Score 2517; DB 4; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.7e-187;
Matches 484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDFTKVSDDTYKLPRLIEKRRD	60
Db	1		
QY	61	RINECIAQLKDPLLPEHLKLTTLGLHLEKAVVLELTLKHLKALTALTEQQHQKIIALONGER	120
Db	61		
QY	121	SLKSPIQSDLDFAHSGFQTCAKEVLOYLSRFESWTPREPRVCVQLINHLHAVATOFLLPTPQ	180
Db	121		
QY	181	LLTQQVPLSKXTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT	240
Db	181		
QY	241	DSYGGEAEARPDPREKKGAGASRVTIKQEPPEGDSAPAKRMKLD SRGGSGGGPGGGAA	300
Db	241		
QY	301	AAAAALLGPDPAANAALLRPDAALLSSLVAFGGGGGAPFPQAAAAAPFCLPFCFLSPSA	360
Db	301		
QY	361	AAAYVQPFLLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAA AAAAAA AAFPCLSSV	420
Db	361		
QY	421	LSPPPEKAGAAAAATLLPHEVAPLGAPHPQHPHGRTHLPZAGPREPGNPESSAQEDPSQPG	480
Db	421		
QY	481	KEAP	484
Db	481		

RESULT 2
AAB70692
ID AAB70692 standard; protein: 482 AA.

The present invention describes a basic helix loop helix (bHLH) type
 transcription factor designated DEC2. DEC2 can be used as a tool in the
 development of drugs for the treatment and prevention of disorders
 involving cell differentiation and proliferation. The present sequence
 represents the specifically claimed human DEC2a protein, as given in the
 present invention
 Sequence 482 AA;

RESULT 3
ABG96308
ID ABG96308 standard; protein: 492 AA.

DEC2 is a basic helix loop helix protein of the DEC family for use in development of drugs for treatment of disorders of cell differentiation and proliferation.

Claim 2: Page 55-59; 83pp; Japanese.

PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
PI Lepley DM, Edinger SR, Burgess CE;
XX
DR WPI; 2003-313242/30.
DR N-PSDB; ACD19338.
XX

PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT and polynucleotides, useful in gene therapy, e.g. for treating or
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
PT stroke or infections.
XX

PS Claim 1; Page 125; 586pp; English.

XX The invention describes a new isolated polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful as therapeutics,
CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This is the amino acid sequence of a novel human NOV protein
XX

SQ Sequence 482 AA;

Query Match 99.2%; Score 2498; DB 6; Length 482;
Best Local Similarity 99.6%; Pred. No. 5.1e-186;
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRD DTKVSDTYKLPRLIEKKRRD 60
DB 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRD DTK--DTYKLPRLIEKKRRD 58
QY 61 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLTKHLKALTALTEQHQKIILQNGER 120
DB 59 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLTKHLKALTALTEQHQKIILQNGER 118
QY 121 SLKSPIQSDLDFAHSGFQTCAKEVLYLSRFESWTPREPCVQLINHLHVAQTFLPTPQ 180
DB 119 SLKSPIQSDLDFAHSGFQTCAKEVLYLSRFESWTPREPCVQLINHLHVAQTFLPTPQ 178
QY 181 LLTQVPLSKGTGAPSAAGSAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT 240
DB 179 LLTQVPLSKGTGAPSAAGSAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT 238
QY 241 DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDSAPAKRMKLD SRGGSGGGPGGAA 300
DB 239 DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDSAPAKRMKLD SRGGSGGGPGGAA 298
QY 301 AAAALLGPDPAALRRPDAALLSSLVAFGGGGGAPFPQPAAPAAAFCLPFCFLSPSA 360
DB 299 AAAALLGPDPAALRRPDAALLSSLVAFGGGGGAPFPQPAAPAAAFCLPFCFLSPSA 358
QY 361 AAAYVQFLDKSGLEKYLYPAAAAAPFPFLYPGIPAPAAAAAFAFPCLSSV 420
DB 359 AAAYVQFLDKSGLEKYLYPAAAAAPFPFLYPGIPAPAAAAAFAFPCLSSV 418
QY 421 LSPPPKAGAAAAATLLPHEVAPLGAPHPQHPHGRTHLPAGPREPGNPRESSAQEDPSQPG 480
DB 419 LSPPPKAGAAAAATLLPHEVAPLGAPHPQHPHGRTHLPAGPREPGNPRESSAQEDPSQPG 478

QY 481 KEAP 484
DB 479 KEAP 482

RESULT 5
ABR58470

ID ABR58470 standard; protein; 482 AA.

XX

AC ABR58470;

XX 07-JUL-2003 (first entry)

XX Human bHLH protein DEC2 from ovarian cancer clone O1668P.

XX Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.

XX Homo sapiens.

XX WO2003029468-A1.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031467.

XX 02-OCT-2001; 2001US-0327135P.

XX 30-MAY-2002; 2002US-0384531P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Mannion J;

XX WPI; 2003-372001/35.

XX New polynucleotide and polypeptide useful for diagnosing and/or treating
PT cancer, particularly ovarian cancer, and as a vaccine.

PS Claim 2; Page 156-157; 169pp; English.

XX The invention relates to a novel isolated polynucleotide. The
CC polynucleotides of the invention have cytostatic activity, and may have a
CC use in gene therapy, and in a vaccine. The composition and methods are
CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.
CC The composition may also be used as a vaccine to prevent cancer. The
CC present sequence is used in the exemplification of the invention

XX SQ Sequence 482 AA;

Query Match 99.2%; Score 2498; DB 6; Length 482;
Best Local Similarity 99.6%; Pred. No. 5.1e-186;
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRD DTKVSDTYKLPRLIEKKRRD 60
DB 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRD DTK--DTYKLPRLIEKKRRD 58
QY 61 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLTKHLKALTALTEQHQKIILQNGER 120
DB 59 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLTKHLKALTALTEQHQKIILQNGER 118
QY 121 SLKSPIQSDLDFAHSGFQTCAKEVLYLSRFESWTPREPCVQLINHLHVAQTFLPTPQ 180
DB 119 SLKSPIQSDLDFAHSGFQTCAKEVLYLSRFESWTPREPCVQLINHLHVAQTFLPTPQ 178
QY 181 LLTQVPLSKGTGAPSAAGSAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT 240
DB 179 LLTQVPLSKGTGAPSAAGSAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT 238
QY 241 DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDSAPAKRMKLD SRGGSGGGPGGAA 300
DB 239 DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDSAPAKRMKLD SRGGSGGGPGGAA 298

DE Human Protein Q9C0J9, SEQ ID NO 467.
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX Homo sapiens.
OS WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 482 AA;

Query Match 99.2%; Score 2498; DB 7; Length 482;
Best Local Similarity 99.6%; Pred. No. 5.1e-186;
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MDEGPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRD DTKVSDTYKLPRLIEKKRRD 60
DB 1 MDEGPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRD DTK--DTYKLPRLIEKKRRD 58
QY 61 RINECTIAQLKDLLPEHLKLTTLGHLEKAVVLETLTKHLKALTALTEQQHQKIALQNGER 120
DB 59 RINECTIAQLKDLLPEHLKLTTLGHLEKAVVLETLTKHLKALTALTEQQHQKIALQNGER 118
QY 121 SLKSPIQSDLD AFHSGFQTC AKEVLQYLSRFESWTPREPRCVQLINHLH AVATQFLPTPQ 180

DB 119 SLKSPIQSDLD AFHSGFQTC AKEVLQYLSRFESWTPREPRCVQLINHLH AVATQFLPTPQ 178
QY 181 LLTQOVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAE LAENDTDT 240
DB 179 LLTQOVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAE LAENDTDT 238
QY 241 DSGYGGEAEARPDREKKGAGASRVTIKQEP PGEDSPAPKRMKLD SRGGGGPGGGAA 300
DB 239 DSGYGGEAEARPDREKKGAGASRVTIKQEP PGEDSPAPKRMKLD SRGGGGPGGGAA 298
QY 301 AAAAALLGPDPA AAAAALLRPDAALLSSLVAFGGGGGAPFPQAAAAAPFCFLSPSA 360
DB 299 AAAAALLGPDPA AAAAALLRPDAALLSSLVAFGGGGGAPFPQAAAAAPFCFLSPSA 358
QY 361 AAAYVQPF LDKSGLEKLYPAAAAAPFLLYPGIPAPAAAAA AAAAAAPFCFLSSV 420
DB 359 AAAYVQPF LDKSGLEKLYPAAAAAPFLLYPGIPAPAAAAA AAAAAAPFCFLSSV 418
QY 421 LSPPEKAGAAAATLLPHEVAPLGAPHPQHPHGRTHLPFAGPREPGNPESSAQEDPSQPG 480
DB 419 LSPPEKAGAAAATLLPHEVAPLGAPHPQHPHGRTHLPFAGPREPGNPESSAQEDPSQPG 478
QY 481 KEAP 484
DB 479 KEAP 482

RESULT 8
ADN05014
ID ADN05014 standard; protein; 482 AA.
XX AC ADN05014;
XX DT 01-JUL-2004 (first entry)
XX DE Antipsoriatic protein sequence #687.
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX PN WO2004028479-A2.
XX PD 08-APR-2004.
XX PF 25-SEP-2003; 2003WO-US030907.
XX PR 25-SEP-2002; 2002US-0414006P.
XX PA (GETH) GENENTECH INC.
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX WPI; 2004-305105/28.
XX N-PSDB; ADN05013.
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX Claim 9; SEQ ID NO 1408; 3069pp; English.
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX
SQ Sequence 482 AA;

Query Match 99.2%; Score 2498; DB 8; Length 482;

Best Local Similarity 99.6%; Pred. No. 5.1e-186;				
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1				
QY	1	MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRD	TKVSDTYKLPRLIEKKRRD 60	
Db	1	MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRD	DTK--DTYKLPRLIEKKRRD 58	
QY	61	RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELT	LKHLKALTALTEQHQKIIALQNGER 120	
Db	59	RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELT	LKHLKALTALTEQHQKIIALQNGER 118	
QY	121	SLKSPIQSDLD AFHSGFQTC AKEVLQYLSRFESWTP	PREPRCVQLINHLHAVATQFLPTPQ 180	
Db	119	SLKSPIQSDLD AFHSGFQTC AKEVLQYLSRFESWTP	PREPRCVQLINHLHAVATQFLPTPQ 178	
QY	181	LLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLE	PLAYCVPVIQRTQPSAELAAENDTDT 240	
Db	179	LLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLE	PLAYCVPVIQRTQPSAELAAENDTDT 238	
QY	241	DSGYGGEAEARPDRKKGKAGASRVTIKQEP	PGEDSPAPKRMKLD	SRGGSGGPGGGAA 300
Db	239	DSGYGGEAEARPDRKKGKAGASRVTIKQEP	PGEDSPAPKRMKLD	SRGGSGGPGGGAA 298
QY	301	AAAAALLGPDPAALRRPDAALLSSLVAF	GGGGAPFPQAAAAAPFCFLSPSA 360	
Db	299	AAAAALLGPDPAALRRPDAALLSSLVAF	GGGGAPFPQAAAAAPFCFLSPSA 358	
QY	361	AAAYVQPFLLKSGLEKLYLPAAAAAPFLLYP	GPAPAAAAA	AAAAAFAFPCLSSV 420
Db	359	AAAYVQPFLLKSGLEKLYLPAAAAAPFLLYP	GPAPAAAAA	AAAAAFAFPCLSSV 418
QY	421	LSPPEKAGAAAATLLPHEVAPLGAPHQPHGR	THLPFAGPREPGNP	ESSAQEDPSQPG 480
Db	419	LSPPEKAGAAAATLLPHEVAPLGAPHQPHGR	THLPFAGPREPGNP	ESSAQEDPSQPG 478
QY	481	KEAP 484		
Db	479	KEAP 482		
RESULT 9				
ABJ04644				
ID	ABJ04644 standard; protein; 767 AA.			
XX	ABJ04644;			
AC	ABJ04644;			
XX	11-OCT-2002 (first entry)			
DT	Protein of NOVX 5 SEQ ID No 10.			
XX	Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipaemic; neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective; tranquiliser; neuroleptic; antidiabetic; antiulcer; antiinflammatory; anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes; metabolic disorder; obesity; infectious disease; Alzheimer's disease; anorexia; neurodegenerative disorder; Parkinson's disorder; obesity; immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease; metabolic syndrome X; wasting disorder; cancer; neurological disorder; epilepsy; stroke; mental disorder; schizophrenia disorders; goiter; vesicular transport; cystic fibrosis; gastrointestinal disorder; diabetes mellitus; ulcerative colitis; AIDS; allergic reaction; multiple sclerosis; rheumatoid arthritis; transgenic animal; gene therapy.			
OS	Unidentified.			
XX	WO200246409-A2.			
PN	13-JUN-2002.			
XX	06-DEC-2001; 2001WO-US046586.			
PD	06-DEC-2000; 2000US-0251660P.			
XX				

PR	12-DEC-2000; 2000US-0255029P.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
----	-------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 524 DSGYGEAEARPDREKKGAGASRVTIKQEPPEGEDSPAPKRMKLDNRGGSGGGPGGAA 583
QY 301 AAAAAALLGPDPAALRRPDAALLSSLVAFGGGGGAPFPQAAAAAPFCLPFCFLSPSA 360
Db 584 AAAAAALLGPDPAALRRPDAALLSSLVAFGGGGGAPFPQAAAAAPFCLPFCFLSPSA 643
QY 361 AAAYVQFLLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAPFCLSSV 420
Db 644 AAAYVQFLLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAPFCLSSV 703
QY 421 LSPPEKAGAAAAATLLPHEVAPLGAPHPQHPHGRTHLPFAGPREPGNPRESSAQEDPSQPG 480
Db 704 LSPPEKAGAAAAATLLPHEVAPLGAPHPQHPHGRTHLPFAGPREPGNPRESSAQEDPSQPG 763
QY 481 KEAP 484
Db 764 KEAP 767

RESULT 10
AAB70694
ID AAB70694 standard; protein; 410 AA.
XX AAB70694;
AC AAB70694;
XX 17-MAY-2001 (first entry)
XX Mouse DEC2a protein sequence SEQ ID NO:14.
DE DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX Mus musculus.
XX WO200114551-A1.
PN 01-MAR-2001.
PD 19-JUN-2000; 2000WO-JP003991.
PF 19-AUG-1999; 99JP-00233286.
PR (CHUS) CHUGAI SEIYAKU KK.
XX Fujimoto K, Shin M, Kato Y;
XX WPI; 2001-202935/20.
DR N-PSDB; AAF74777.
XX DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation.
XX Claim 2; Page 77-80; 83pp; Japanese.
XX The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in the
CC development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC represents the specifically claimed mouse DEC2a protein, as given in the
CC present invention
XX Sequence 410 AA;

Query Match 64.9%; Score 1634; DB 4; Length 410;
Best Local Similarity 72.0%; Pred. No. 9.3e-119;
Matches 350; Conservative 21; Mismatches 37; Indels 78; Gaps 16;
QY 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDITKVSDDTYKLPRLIEKKRD 60
Db 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSLKRDITK--DTYKLPRLIEKKRD 58

QY 61 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKHLKALTALTEQQHQKIALLQNGER 120
Db 59 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKHLKALTALTEQQHQKIALLQNGER 118
QY 121 SLKSPIQSDLDAPHSGFOTCAKEVLQYLSRFESWTPREPCVQLINHLHAVATQFLPTPQ 180
Db 119 SLKSPVQADLDAPHSGFOTCAKEVLQYLSRFESWTPREPCVQLINHLHAVAT-----Q 172
QY 181 LLTQQVPLSKGTG-APSAAGSAAAPCLERAGQKLEPLAYCVPIQRTQPSAELAAENDTD 239
Db 173 LLTPQVPSGRSGRAPCSAGAAAAAGPER-----VARCVPIQRTQPGTE--PEHDT 223
QY 240 TDSGYGGEAEARPDREKKGAGASRVTIKQEPPEGEDSPAPKRMKLDNRGGSGGGPGGGA 299
Db 224 TDSGYGGEAE-----QGRAAVKQEPFGDSSPAPKRPKLEARG----- 260
QY 300 AAAAAALLGPDPAALRRPDAALLSSLVAFGGGGGAPFPQAAAAAPFCLPFCFLSPS 359
Db 261 -----ALLGPEP-----ALLGSLVAL--GGGAPFAQP--AAAPFCLPFCFLSPS 300
QY 360 AAAYVQFLLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAPFCLSS 419
Db 301 -AAAYVQFLLDKSGLEKYLYP-AAAAAPFLLYPGIP-----AAAAAAPFCLSS 350
QY 420 VLSPPPEKAGA-AAATLLPHEVAPLGAPHPQHPHGRTHLPFAGPREPGNPRESSAQEDPSQ 478
Db 351 VLSPPPEKAGATAGAPFLAHEVAPPGPLRPQHAHSRTHLPRA-----VNPESS-QEDATQ 404
QY 479 PGKEAP 484
Db 405 PAKDAP 410

RESULT 11
ADE54660
ID ADE54660 standard; protein; 410 AA.
XX ADE54660;
AC ADE54660;
XX 29-JAN-2004 (first entry)
DT Rat Protein O35779, SEQ ID NO 465.
DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX Rattus norvegicus.
XX WO2003016475-A2.
PN 27-FEB-2003.
PD 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
XX GENBANK; O35779.

New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX Claim 1; Page; 1017pp; English.
PS The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
CC


```
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This is the amino acid sequence of a novel human NOV protein
XX
SQ Sequence 203 AA;

Query Match          36.0%; Score 905; DB 6; Length 203;
Best Local Similarity 88.2%; Pred. No. 2.5e-62;
Matches 179; Conservative 2; Mismatches 12; Indels 10; Gaps 2;

QY 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDYKVSQDTYKLPRLIEKKRRD 60
Db 3 MDEGIPHLQERQLPEHRDFIGLDYSSLYMCKPKRSMKRDYK--DTYKLPRLIEKKRRD 60

QY 61 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTALTEQQHQKIIALQNGER 120
Db 61 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTALTEQQHQKIIALQNGER 120

QY 121 SLKSPIQSDLDAFHSGFQTCAKEVLYLSRFESWTPREPRCVQLINHLHAVATQFLP-- 177
Db 121 SLKSPIQSDLDAFHSGFQTCAKEVLYLSRFEGWTPREPRCVQLINHLHAVATQFLPFAG 180

QY 178 -----TPQLLTQQVPLSKGTGAP 195
Db 181 PREPGNPRESSAQEDPSQPGKEAP 203

RESULT 13
AAU16188
ID AAU16188 standard; protein; 165 AA.
XX
AC AAU16188;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human novel secreted protein, Seq ID 1141.
XX
KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
```

PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488783/53.
DR N-PSDB; AAS26175.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
PS Claim 11; SEQ ID NO 1141; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac

CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 33.2%; Score 835; DB 4; Length 165;
Best Local Similarity 98.8%; Pred. No. 5.5e-57;
Matches 165; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 37 KRDDTKVSDTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLK 96
Db 1 KRDDTK--DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLK 58
QY 97 HLKALTALTEQHQKIIALQNGERSLKSPISQSDLDFAHSGFQCAKEVLQYLSRFESWTP 156
Db 59 HLKALTALTEQHQKIIALQNGERSLKSPISQSDLDFAHSGFQCAKEVLQYLSRFESWTP 118
QY 157 REPRCVQLINHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAA 203
Db 119 REPRCVQLINHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAA 165

RESULT 14
ABU55257
ID ABU55257 standard; protein; 165 AA.
XX
AC ABU55257;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polypeptide #344.
XX
KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PN US2002132753-A1.
XX
PD 19-SEP-2002.
XX
PF 17-JAN-2001; 2001US-00764864.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.

PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.
N-PSDB; AAS26590.

New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 11; SEQ ID NO 1556; 980pp; English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence represents a novel secreted protein of the invention. Note: The sequence data for this patent did not form part of the printed

Query Match 29.8%; Score 750; DB 4; Length 150;
Best Local Similarity 97.4%; Pred. No. 2.1e-50;
Matches 148; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY	37	KRDDTKVSDTYKLP	PHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLK	96
Db	1	KRDDTK--DTYK	LXHLRIEKKRRDRINECIAQLKDLLPEHLKLTXLGHLEKAVVLELTLK	58
QY	97	HLKALTALTEQOQHQ	KIIALQNGERSLKSPIQSDLDAFHSGFQTC	156
Db	59	HLKALTALTEQOQHQ	KIIALQNGERSLKSPIQSDLDAFHSGFQTC	118
QY	157	REPRCVQLINHLH	AVATQFLPTPQLLTQQVPL	188
Db	119	REPRCVQLINHLH	AVATQFLPTPQLLTQQVPL	150

Search completed: December 15, 2004, 19:52:24
Job time : 163 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 17, 2004, 03:13:29 ; Search time 688.631 Seconds
(without alignments)
11518.321 Million cell updates/sec

Title: US-10-078-650-11
Perfect score: 1511
Sequence: 1 catggacgaaggaattcctc.....gtgagaagttaataaccct 1511

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size : 0
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1511	100.0	1511	4	Aaf74776 Human DEC
2	1378	91.2	3521	10	Adc30675 Human nov
3	1378	91.2	3641	4	Aaf74767 Human DEC
4	1378	91.2	3641	6	Abs76400 cDNA enco
5	1378	91.2	3641	9	Acc72416 Human ova
6	1378	91.2	3641	12	Adn05013 Antipsori
7	1339	88.6	1471	10	Acd19338 cDNA enco
8	1323	87.6	2304	6	Abt05457 DNA of NO
9	552	36.5	909	10	Abx74439 Human cdn
10	476	31.5	498	4	Aas26175 Human cdn
11	476	31.5	498	8	Abx73516 Human nov
12	366	24.2	454	4	Aas26590 Human cdn
13	366	24.2	454	8	Abx73931 Human nov
14	342	22.6	628	10	Acd19339 cDNA enco
15	274	18.1	285	4	Aas24787 Human ova
16	274	18.1	397	5	Aah83421 Human ova
17	128	8.5	502	3	Aac04136 Human sec
18	128	8.5	896	10	Adc32475 Human nov
19	54	3.6	1421	4	Aaf74777 Mouse DEC
20	46	3.0	300	3	Aaa01182 Human col
21	32	2.1	8147	6	Abs73313 DNA enco

C	22	32	2.1	8147	10	ADK11302	Adk11302 Human CRE
	23	31	2.1	1679	8	ABZ78139	Abz78139 Human can
C	24	31	2.1	3060	2	AAX39667	Aax39667 Renal can
C	25	31	2.1	4725	5	AAS76375	Aas76375 DNA enco
C	26	31	2.1	4739	6	ABL63316	AbL63316 Breast ca
C	27	31	2.1	4739	8	ACC00031	Acc00031 Human Bcr
C	28	31	2.1	4739	10	ADG32836	Adg32836 Human DNA
	29	31	2.1	4756	5	AAS85030	Aas85030 DNA enco
C	30	31	2.1	4775	5	AAS76377	Aas76377 DNA enco
C	31	31	2.1	5795	5	AAS85031	Aas85031 DNA enco
C	32	30	2.0	884	12	ADE28379	Ade28379 Human KPP
C	33	30	2.0	3544	5	AAS81565	Aas81565 DNA enco
C	34	30	2.0	6335	6	ABK83722	Abk83722 Human cdn
C	35	30	2.0	6335	11	ADN95256	Adn95256 Human BEC
C	36	29	1.9	186	12	ADF77133	Adf77133 M08-TAG C
C	37	29	1.9	201	12	ADF77129	Adf77129 E25-TAG C
	38	29	1.9	526	12	ACH78213	Ach78213 Human gen
C	39	29	1.9	529	12	ACH77958	Ach77958 Human gen
	40	29	1.9	546	6	ABQ41782	Abq41782 Oligonucl
C	41	29	1.9	546	6	ABQ41783	Abq41783 Oligonucl
C	42	29	1.9	734	5	ADL46045	Adl46045 Human ova
C	43	29	1.9	797	6	ABK34450	Abk34450 Human cdn
	44	29	1.9	1130	6	ABK92146	Abk92146 Prostate
	45	29	1.9	1130	12	ADP13302	Adp13302 Renal cel

ALIGNMENTS

RESULT 1
AAF74776
ID AAF74776 standard; cDNA; 1511 BP.
XX
AC AAF74776;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human DEC2b encoding cDNA sequence SEQ ID NO:11.
XX
KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
XW basic helix loop helix protein; cell differentiation; proliferation; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 2..1456
FT /*tag= a
FT /product= "DEC2b"
FT /note= "bHLH type transcription factor"

XX
PN WO200114551-A1.
XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP003991.
XX
PR 19-AUG-1999; 99JP-00233286.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Fujimoto K, Shin M, Kato Y;
XX
DR WPI; 2001-202935/20.
DR P-PSDB; AAB70693.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
XX and proliferation.
PS Claim 1; Page 63-68; 83pp; Japanese.
XX
CC The present invention describes a basic helix loop helix (bHLH) type
transcription factor designated DEC2. DEC2 can be used as a tool in the

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR P-PSDB; ADC31646.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 757; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3521 BP; 937 A; 902 C; 851 G; 831 T; 0 U; 0 Other;
Query Match 91.2%; Score 1378; DB 10; Length 3521;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAAATGAA 193
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
377 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAAATGAA 436
QY 194 TGCATTGCTCAGCTGAAAGATTCTAGCTGAACATCTGAATTGACAACTCTGGGACAT 253
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
437 TGCATTGCTCAGCTGAAAGATTCTAGCTGAACATCTGAATTGACAACTCTGGGACAT 496
QY 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAACACTTAAAGCTTTAACCGCCTTA 313
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
497 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAACACTTAAAGCTTTAACCGCCTTA 556
QY 314 ACCGAGCAACACCATCAGAAGATAATTGCTTTACAGAATGGGAGCGGATCTCTGAATCG 373
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
557 ACCGAGCAACACCATCAGAAGATAATTGCTTTACAGAATGGGAGCGGATCTCTGAATCG 616
QY 374 CCCATTGACGCTTGGATGCGTTCACACTCGGGATTTCAAACATGCGCCAAAGAGTC 433
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
617 CCCATTGACGCTTGGATGCGTTCACACTCGGGATTTCAAACATGCGCCAAAGAGTC 676

QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGGTGTGTCCAGCTG 493
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
677 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGGTGTGTCCAGCTG 736
QY 494 ATCAACCACTTGACAGCCGCTGGCCACCCAGTTTCTTGCCCAACCCTGCTGACTCAA 553
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
737 ATCAACCACTTGACAGCCGCTGGCCACCCAGTTTCTTGCCCAACCCTGCTGACTCAA 796
QY 554 CAGGTCCCTCTGAGCAAAAGGCACCGGCGCTCCCTCGGCCCGCGGTCGCGGCCGCC 613
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
797 CAGGTCCCTCTGAGCAAAAGGCACCGGCGCTCCCTCGGCCCGCGGTCGCGGCCGCC 856
QY 614 TGCCTGGAGCGCGGGGCGAGAAAGCTGGAGCCCTCTGCGCTACTGCTGCCGTATCCAG 673
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
857 TGCCTGGAGCGCGGGGCGAGAAAGCTGGAGCCCTCTGCGCTACTGCTGCCGTATCCAG 916
QY 674 CGGACTCAGCCAGCCGCGAGCTCGCCGCGGAGAACGACACGACACCGACGCGGCTAC 733
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
917 CGGACTCAGCCAGCCGCGAGCTCGCCGCGGAGAACGACACGACACCGACGCGGCTAC 976
QY 734 GGCGGCGAAGCCGAGGCCCGCGGACCGCGAGAAAGGCAAGGCGGGGGCGGAGCCGC 793
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
977 GGCGGCGAAGCCGAGGCCCGCGGACCGCGAGAAAGGCAAGGCGGGGGCGGAGCCGC 1036
QY 794 GTCACCATCAAGCAGGAGCTCCCGGGGAGGACTCGCCGGCGCCCAAGAGGATGAAGCTG 853
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1037 GTCACCATCAAGCAGGAGCTCCCGGGGAGGACTCGCCGGCGCCCAAGAGGATGAAGCTG 1096
QY 854 GATTCCCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 913
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1097 GATTCCCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1156
QY 914 GCGTTCTGGGCGCGGAGCTTCCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 973
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1157 GCGTTCTGGGCGCGGAGCTTCCCGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCG 1216
QY 974 CTCAGCTCGTGGTGGCGTTTCGGCGGAGGCGGAGGCGGCGGCGGCGGCGGCGGCG 1033
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1217 CTCAGCTCGTGGTGGCGTTTCGGCGGAGGCGGAGGCGGCGGCGGCGGCGGCGGCG 1276
QY 1034 GCCGCGGCGGCGGCTTCTGCTGCGGCTTCTGCTTCTCTGCGCTTCTGAGCTGCGGCTAC 1093
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1277 GCCGCGGCGGCGGCTTCTGCTGCGGCTTCTGCTTCTCTGCGCTTCTGAGCTGCGGCTAC 1336
QY 1094 GTGAGCCCTTCTGAGCAAGAGCGGCGGCTGGAGAGTATCTGTACCGCGGCGGCTGCC 1153
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1337 GTGAGCCCTTCTGAGCAAGAGCGGCGGCTGGAGAGTATCTGTACCGCGGCGGCTGCC 1396
QY 1154 GCGCGGCTTCTGCTATACCGCGGCTATACCGCGGCTATACCGCGGCTATACCGCGG 1213
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1397 GCGCGGCTTCTGCTATACCGCGGCTATACCGCGGCTATACCGCGGCTATACCGCGG 1456
QY 1214 GCGCGGCTTCTGCGCGGCGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1457 GCGCGGCTTCTGCGCGGCGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1516
QY 1274 CCGGAGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1333
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1517 CCGGAGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1576
QY 1334 GCGCGGCTATACCGCGGCTATACCGCGGCTATACCGCGGCTATACCGCGGCTATAC 1393
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1577 GCGCGGCTATACCGCGGCTATACCGCGGCTATACCGCGGCTATACCGCGGCTATAC 1636
QY 1394 CCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGAGCCAGGAAGAGCTCCC 1453
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1637 CCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGAGCCAGGAAGAGCTCCC 1696
QY 1454 TGAATCCTTGGTCCGAAAGGACCGGAGGTTCAAGCAGAGTGAGAGTAAATACCTT 1511
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1697 TGAATCCTTGGTCCGAAAGGACCGGAGGTTCAAGCAGAGTGAGAGTAAATACCTT 1754

```
RESULT 3
AAF74767
ID AAF74767 standard; cDNA; 3641 BP.
XX AC
XX AAF74767;
XX DT
XX 17-MAY-2001 (first entry)
XX DE
XX Human DEC2a encoding cDNA sequence SEQ ID NO:1.
XX KW
XX DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
XX KW basic helix loop helix protein; cell differentiation; proliferation; ss.
XX OS
XX Homo sapiens.
XX FH
XX Key Location/Qualifiers
XX CDS 135..1583
XX FT /*tag= a
XX FT /product= "DEC2a"
XX FT /note= "bHLH type transcription factor"
XX PN
XX WO200114551-A1.
XX XX
XX PD
XX 01-MAR-2001.
XX PF
XX 19-JUN-2000; 2000WO-JP003991.
XX PR
XX 19-AUG-1999; 99JP-00233286.
XX PA
XX (CHUS ) CHUGAI SEIYAKU KK.
XX XX
XX PI
XX Fujimoto K, Shin M, Kato Y;
XX XX
XX WPI; 2001-202935/20.
XX DR
XX P-PSDB; AAB70692.
XX PT
XX DEC2 is a basic helix loop helix protein of the DEC family for use in
XX PT development of drugs for treatment of disorders of cell differentiation
XX PT and proliferation.
XX XX
XX PS
XX Claim 1; Page 48-55; 83pp; Japanese.
XX CC
XX The present invention describes a basic helix loop helix (bHLH) type
XX CC transcription factor designated DEC2. DEC2 can be used as a tool in the
XX CC development of drugs for the treatment and prevention of disorders
XX CC involving cell differentiation and proliferation. The present sequence
XX CC encodes the specifically claimed human DEC2a protein, as given in the
XX CC present invention
XX SQ
XX Sequence 3641 BP; 994 A; 922 C; 836 G; 889 T; 0 U; 0 Other;
Query Match 91.2%; Score 1378; DB 4; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 134 GATACCTACAAATTACCGCAGAGTAATAGAAAAGAAAAGAGACCGGAATTAATGAA 193
Db 261 GATACCTACAAATTACCGCAGAGTAATAGAAAAGAAAAGAGACCGGAATTAATGAA 320
QY 194 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGGACAT 253
Db 321 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGGACAT 380
QY 254 CTGGAGAAAGCTGTAGTCTTGAATTAACCTTGAACACTTAAAGCTTTAACCGCCTTA 313
Db 381 CTGGAGAAAGCTGTAGTCTTGAATTAACCTTGAACACTTAAAGCTTTAACCGCCTTA 440
QY 314 ACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 373
Db 441 ACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 500
QY 374 CCCATTGCTCCGACTTGGATGGTCCACTCGGGATTTCAAACATGCGCCAAAGAGTC 433
```

```
Db 501 CCATTTCAGTCCGACTTGGATGCGTTCCACTCGGATTTCAAACATGCGCAAAGAGTC 560
QY 434 TTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGGAGCCCGGTGTGTCCAGCTG 493
Db 561 TTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGGAGCCCGGTGTGTCCAGCTG 620
QY 494 ATCAACCACTTGACGCGCGTGGCCACCCAGTTCTTGCCCAACCCCGCAGCTGTGACTCAA 553
Db 621 ATCAACCACTTGACGCGCGTGGCCACCCAGTTCTTGCCCAACCCCGCAGCTGTGACTCAA 680
QY 554 CAGGTCCCTCTGAGCAAAAGCACCAGCGCTCCCTCGGCCCGCGGTCCGGCGCGCCCCC 613
Db 681 CAGGTCCCTCTGAGCAAAAGCACCAGCGCTCCCTCGGCCCGCGGTCCGGCGCGCCCCC 740
QY 614 TGCCTGGAGCGCGCGGCGAGAAAGCTGGAGCCCTCGCCTACTGCTGCCGTCTCATCCAG 673
Db 741 TGCCTGGAGCGCGCGGCGAGAAAGCTGGAGCCCTCGCCTACTGCGTGCCTCATCCAG 800
QY 674 CGGACTCAGCCAGCGCGGAGCTCGCCCGCGAGAACACGACACCGACAGCGGCTAC 733
Db 801 CGGACTCAGCCAGCGCGGAGCTCGCCCGCGAGAACACGACACCGACAGCGGCTAC 860
QY 734 GCGCGGAAGCCGAGGCCCGCGGACCCGCGAGAAAGGCAAGGCGCGGCGCGGCGCGCCG 793
Db 861 GCGCGGAAGCCGAGGCCCGCGGACCCGCGAGAAAGGCAAGGCGCGGCGCGGCGCGCCG 920
QY 794 GTCACCATCAAGCAGGAGCTCCCGGGGAGGACTCGCCGGCGCCCAAGAGGATGAAGCTG 853
Db 921 GTCACCATCAAGCAGGAGCTCCCGGGGAGGACTCGCCGGCGCCCAAGAGGATGAAGCTG 980
QY 854 GATCCCGCGCGCGGCGAGCGCGGCGCGCCCGGGGGCGCGCGCGCGCGCGCGCGCGCCG 913
Db 981 GATCCCGCGCGCGGCGAGCGCGCGCGCGCCCGGGGGCGCGCGCGCGCGCGCGCGCGCCG 1040
QY 914 GCGCTTCTGGGCGCCGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCCG 973
Db 1041 GCGCTTCTGGGCGCCGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCCG 1100
QY 974 CTCAGCTCGTGGTGGCTTCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033
Db 1101 CTCAGCTCGTGGTGGCTTCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1160
QY 1034 GCGCGGCGCCCTTCTGCTGCTGCCCTTCTGCTTCTGCTTCTGCTTCTGCTGCTGCTTCT 1093
Db 1161 GCGCGGCGCCCTTCTGCTGCTGCCCTTCTGCTTCTGCTTCTGCTTCTGCTGCTGCTTCT 1220
QY 1094 GTGAGCCCTTCTCTGGACAAGAGCGGCGCTTGAGAGAGTATCTGTACCCGGCGCGGCTGCC 1153
Db 1221 GTGAGCCCTTCTCTGGACAAGAGCGGCGCTTGAGAGAGTATCTGTACCCGGCGCGGCTGCC 1280
QY 1154 GCGCGGTTCCCGCTGTATATACCCCGGCGATCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1213
Db 1281 GCGCGGTTCCCGCTGTATATACCCCGGCGATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1340
QY 1214 GCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1273
Db 1341 GCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1400
QY 1274 CCGAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1333
Db 1401 CCGAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1460
QY 1334 GCGCGGCAACCCCGAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1393
Db 1461 GCGCGGCAACCCCGAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1520
QY 1394 CCGGGGAACCCGGAGAGCTCTGCTCAGGAAGATCCCTGCAGCCAGGAAGAGAGCTCCC 1453
Db 1521 CCGGGGAACCCGGAGAGCTCTGCTCAGGAAGATCCCTGCAGCCAGGAAGAGAGCTCCC 1580
QY 1454 TGAATCCTTGGTCCCGAAGGACGGAGGTTCAAGCAGAGTGAGAGTTAAATACCCCT 1511
```


Db	1281	GCCTCGTTCCTGCTATACCCCGGCATCCCGCCCGCGGCAGCCGCGCGCC	1340
Qy	1214	GCGCGCGCTGCGCGCGCGCGCGCGTTCCTCGCTGCTCGCTGCTCGCGCGCT	1273
Db	1341	GCGCGCGCTGCGCGCGCGCGCGCTTCCTCGCTGCTCGCTGCTCGCGCGCT	1400
Qy	1274	CCGAGAGCGCGCGCGCGCGCGCTTCCTCGCGCACGAGGTGGCGCGCTTGGG	1333
Db	1401	CCGAGAGCGCGCGCGCGCGCGCTTCCTCGCGCACGAGGTGGCGCGCTTGGG	1460
Qy	1334	GCGCGCACCCCGCGCGCGCGCGCGCTTCCTCGCGCACGAGGTGGCGCGCTTGGG	1393
Db	1461	GCGCGCACCCCGCGCGCGCGCGCGCTTCCTCGCGCACGAGGTGGCGCGCTTGGG	1520
Qy	1394	CCGCGGACCCCGGAGAGCTCTGCTCAGGAAGATCCCTCGCGCACGAGGTGGCGCGCTTGGG	1453
Db	1521	CCGCGGACCCCGGAGAGCTCTGCTCAGGAAGATCCCTCGCGCACGAGGTGGCGCGCTTGGG	1580
Qy	1454	TGAATCCTTGGCTCCGAGGAGCGGAGGTTCAAGCAGAGTGAGAGTTAAATACCT	1511
Db	1581	TGAATCCTTGGCTCCGAGGAGCGGAGGTTCAAGCAGAGTGAGAGTTAAATACCT	1638
RESULT 7			
ACD19338			
ID	ACD19338 standard; cDNA; 1471 BP.		
AC	ACD19338;		
XX			
DT	25-AUG-2003 (first entry)		
XX			
DE	cDNA encoding novel human protein #18.		
XX			
KW	Human; NOV; gene therapy; endocrine related disease; diabetes;		
KW	metabolism-related disease; obesity; central nervous system disorder;		
KW	Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;		
KW	schizophrenia; depression; autoimmune disorder; inflammatory disorder;		
KW	psoriasis; allergy; lupus erythematosus; asthma; cancer;		
KW	inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;		
KW	colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;		
KW	prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;		
KW	lung disease; emphysema; obstructive pulmonary disease; haemophilia;		
KW	stroke; infection; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2003023002-A2.		
XX			
PD	20-MAR-2003.		
XX			
PF	09-SEP-2002; 2002WO-US028539.		
XX			
PR	07-SEP-2001; 2001US-0318120P.		
PR	07-SEP-2001; 2001US-0318130P.		
PR	10-SEP-2001; 2001US-0318430P.		
PR	17-SEP-2001; 2001US-0322636P.		
PR	17-SEP-2001; 2001US-0322781P.		
PR	17-SEP-2001; 2001US-0322816P.		
PR	17-SEP-2001; 2001US-0322817P.		
PR	19-SEP-2001; 2001US-0323519P.		
PR	20-SEP-2001; 2001US-0323631P.		
PR	20-SEP-2001; 2001US-0323636P.		
PR	25-SEP-2001; 2001US-0324969P.		
PR	25-SEP-2001; 2001US-0325091P.		
PR	26-SEP-2001; 2001US-0324990P.		
PR	17-APR-2002; 2002US-0373212P.		
PR	06-SEP-2002; 2002US-00236177.		
XX			
PA	(CURA-) CURAGEN CORP.		
XX			
PI	Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;		
PI	Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;		
PI	Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;		

PI	Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;	
PI	Lepley DM, Edinger SR, Burgess CE;	
XX		
DR	WPI; 2003-313242/30.	
DR	P-PSDB; ABO14645.	
XX		
PT	New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)	
PT	and polynucleotides, useful in gene therapy, e.g. for treating or	
PT	preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,	
PT	stroke or infections.	
XX		
PS	Claim 20; Page 125; 586pp; English.	
XX		
CC	The invention describes a new isolated polypeptide (NOVX). The NOVX	
CC	polypeptide, nucleic acid and antibody are useful as therapeutics,	
CC	particularly in the manufacture of a medicament for treating a syndrome	
CC	associated with a human disease, which includes a pathology associated	
CC	with NOVX polypeptide. The DNA encoding the protein is useful in gene	
CC	therapy for treating the disease or condition. In particular, the NOVX	
CC	polypeptide or polynucleotide is useful for treating endocrine/	
CC	metabolism-related diseases (e.g. obesity or diabetes), central nervous	
CC	system disorders (e.g. Alzheimer's disease, Parkinson's disease,	
CC	epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune	
CC	and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,	
CC	asthma, inflammatory bowel disease, rheumatoid arthritis or	
CC	osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,	
CC	prostate or brain cancers, or melanoma), liver diseases (e.g. liver	
CC	cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),	
CC	haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).	
CC	These are also useful in developing powerful assay system for functional	
CC	analysis of various human disorders, as well as in diagnostic	
CC	applications, and for monitoring the effects of drugs during clinical	
CC	trials. This sequence encodes a novel human NOV protein	
XX		
SQ	Sequence 1471 BP; 288 A; 515 C; 436 G; 232 T; 0 U; 0 Other;	
Query Match 88.6%; Score 1339; DB 10; Length 1471;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 1339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	134	GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAATGAA 193
Db	133	GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAATGAA 192
Qy	194	TGCATTGCTCAGCTGAAAGATTACTGCTGAACATCTGAAATTGACAACTCTGGGACAT 253
Db	193	TGCATTGCTCAGCTGAAAGATTACTGCTGAACATCTGAAATTGACAACTCTGGGACAT 252
Qy	254	CTGAGAAAGCTGTAGTCTTGGAATTAACCTTGAAACACTTAAAGCTTTAACCGCCTTA 313
Db	253	CTGAGAAAGCTGTAGTCTTGGAATTAACCTTGAAACACTTAAAGCTTTAACCGCCTTA 312
Qy	314	ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 373
Db	313	ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 372
Qy	374	CCCATTCAGTCCGACTTGGATGCGTTCCACTCGGATTTCAAACATGCGCCAAAGAGTC 433
Db	373	CCCATTCAGTCCGACTTGGATGCGTTCCACTCGGATTTCAAACATGCGCCAAAGAGTC 432
Qy	434	TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCCGGTGTCTCCAGCTG 493
Db	433	TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCCGGTGTCTCCAGCTG 492
Qy	494	ATCAACCACTTGACCGCGTGGCCACCCAGTTCTTGCCCAACCCCGCAGCTGTGACTCAA 553
Db	493	ATCAACCACTTGACCGCGTGGCCACCCAGTTCTTGCCCAACCCCGCAGCTGTGACTCAA 552
Qy	554	CAGTCCCTCTGAGCAAGGACCGCGGCTCCCTCGGCGCGCGGTCCGCGCGCGCCCC 613
Db	553	CAGTCCCTCTGAGCAAGGACCGCGGCTCCCTCGGCGCGCGGTCCGCGCGCGCCCC 612
Qy	614	TGCTGGAGCGCGGGGAGCAAGCTGGAGCCCTCGCCTACTGCGTGCCTCATCCAG 673

Db 613 TGCCTGGAGCGCGCGGCAGAAAGCTGGAGCCCTGCCTACTGCGTGCCGTCATCCAG 672
QY 674 CGGACTCAGCCCGAGCGCGGAGCTCGCCGCGGAGAACGACACGACACCCAGCGGCTAC 733
Db 673 CGGACTCAGCCCGAGCGCGGAGCTCGCCGCGGAGAACGACACGACACCCAGCGGCTAC 732
QY 734 GCGCGGGAAGCCGAGCGCGCGGCGGACCGCGGAGAAAGCAAGCGCGGGGCGAGCCGC 793
Db 733 GCGCGGGAAGCCGAGCGCGCGGCGGACCGCGGAGAAAGCAAGCGCGGGGCGAGCCGC 792
QY 794 GTCACCATCAAGCAGGAGCCTCCCGGGAGGAGTCCCGGGCGGCTCGCGGCGCCCAAGAGGATGAAGCTG 853
Db 793 GTCACCATCAAGCAGGAGCCTCCCGGGAGGAGTCCCGGGCGGCTCGCGGCGCCCAAGAGGATGAAGCTG 852
QY 854 GATTCCCGCGCGCGGCGGAGCGCGGCGGCGCGCGCGGCGGCGCGCGCGCGCGCGCG 913
Db 853 GATTCCCGCGCGCGGCGGAGCGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 912
QY 914 GCGCTTCTGGGGCCCGACCCCTGCGCGCGCGCGCGCGCGCTGCTGAGACCGCGCGCCCTG 973
Db 913 GCGCTTCTGGGGCCCGACCCCTGCGCGCGCGCGCGCGCTGCTGAGACCGCGCGCCCTG 972
QY 974 CTCAGCTCGTGGTGGCGTTCCGCGGAGCGCGCGCGCGCGCTTCCCGAGCCCGCGGCC 1033
Db 973 CTCAGCTCGTGGTGGCGTTCCGCGGAGCGCGCGCGCGCGCTTCCCGAGCCCGCGGCC 1032
QY 1034 GCGCGGCGCCCTTCTGCTGCCCTTCTGCTTCTCTCGCTTCTGAGAGTCCCGCTAC 1093
Db 1033 GCGCGGCGCCCTTCTGCTGCCCTTCTGCTTCTCTCGCTTCTGAGAGTCCCGCTAC 1092
QY 1094 GTGAGCCCTTCTGAGAACAGAGCGGCTTGAGAGAGTATCTGATCCCGCGCGGCTGCC 1153
Db 1093 GTGAGCCCTTCTGAGAACAGAGCGGCTTGAGAGAGTATCTGATCCCGCGCGGCTGCC 1152
QY 1154 GCCCGGTTCCCGCTGCTATACCCCGGCGATCCCGCGCGCGCGCGCGCGCGCGCGCC 1213
Db 1153 GCCCGGTTCCCGCTGCTATACCCCGGCGATCCCGCGCGCGCGCGCGCGCGCGCGCC 1212
QY 1214 GCGCGCGCTGCGCGCGCGCGCGGTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
Db 1213 GCGCGCGCTGCGCGCGCGCGCGGTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
QY 1274 CCGGAGAGCGCGCGCGCGCGCGGACCTCTGCGCGACGAGTGCGCGCCCTTGGG 1333
Db 1273 CCGGAGAGCGCGCGCGCGCGCGGACCTCTGCGCGACGAGTGCGCGCCCTTGGG 1332
QY 1334 GCGCGGCAACCCAGCACCCGACGCGCGGACCCACCTGCGCGCGCGCGCGCGCGCGGAG 1393
Db 1333 GCGCGGCAACCCAGCACCCGACGCGCGGACCCACCTGCGCGCGCGCGCGCGCGGAG 1392
QY 1394 CCGGGGAACCCGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGGAGCTCCC 1453
Db 1393 CCGGGGAACCCGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGGAGCTCCC 1452
QY 1454 TGAATCCTTGCCTCCCGAA 1472
Db 1453 TGAATCCTTGCCTCCCGAA 1471

RESULT 8
ABT05457
ID ABT05457 standard; DNA; 2304 BP.
XX
AC ABT05457;
XX
DT 11-OCT-2002 (first entry)
XX
DE DNA of NOVX 5 SEQ ID No 9.
XX

KW Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipaeamic;
KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
KW tranquiliser; neuroleptic; antidiabetic; antiulcer; antiinflammatory;

KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
KW vesicular transport; cystic fibrosis; gastrointestinal disorder;
KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
XX gene therapy; gene; ds.
OS Unidentified.
XX WO200246409-A2.
FN 13-JUN-2002.
XX 06-DEC-2001; 2001WO-US046586.
PF 06-DEC-2000; 2000US-0251660P.
XX 12-DEC-2000; 2000US-0255029P.
PR 08-JAN-2001; 2001US-0260326P.
PR 24-JAN-2001; 2001US-0263800P.
PR 20-FEB-2001; 2001US-0269942P.
PR 24-APR-2001; 2001US-0286183P.
PR 20-AUG-2001; 2001US-0313627P.
PR 12-SEP-2001; 2001US-0318712P.
XX (CURA-) CURAGEN CORP.
PA Guo X, Li L, Patturajan M, Shimkets RA, Casman SJ, Malyankar UM;
XX Tchernev VT, Vernet CAM, Spytek KA, Shenoy SG, Alsobrook JP;
PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA, Boldog FL;
PI Colman SD, Eisen AJ, Liu X, Padigar M, Spaderna SK, Zerhusen BD;
XX WPI; 2002-547774/58.
DR P-PSDB; ABJ04644.
XX Novel isolated polypeptide, designated NOVX, useful for treating or
PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT metabolic, neurodegenerative, immune and hematopoietic disorders.
XX Claim 9; Page 43-44; 421pp; English.
PS The invention relates to an isolated polypeptide, designated NOVX,
XX comprising a sequence fully defined in the specification. The isolated
CC protein, its encoding polynucleotide or an antibody created from the
CC protein is useful in the manufacture of a medicament for treating a
CC syndrome associated with a human disease, preferably a NOVX-associated
CC disorder, or for treating or preventing a NOVX-associated disorder in a
CC subject, preferably human. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are also useful
CC for treating or preventing metabolic disorders, diabetes, obesity,
CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
CC disease, Parkinson's disorder, immune disorders, haematopoietic
CC disorders, and various dyslipidaemias, metabolic disturbances associated
CC with obesity, the metabolic syndrome X, wasting disorders associated with
CC chronic diseases, and cancer. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are useful for
CC treating or preventing neurological disorders such as epilepsy, stroke,
CC mental disorders including mood, anxiety, schizophrenic disorders,
CC disorders of vesicular transport such as cystic fibrosis, diabetes
CC mellitus, goiter, gastrointestinal disorders including ulcerative
CC colitis, other conditions associated with abnormal vesicle trafficking
CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
CC arthritis. A cell comprising the vector of the invention is useful for
CC producing non-human transgenic animals. The polynucleotide of the
CC invention can be used to treat disorders by gene therapy. This
CC polynucleotide sequence represents the DNA encoding one of the isolated
CC NOVX proteins of the invention
XX Sequence 2304 BP; 468 A; 784 C; 689 G; 363 T; 0 U; 0 Other;
SQ

CC	represents a human cDNA sequence up-regulated in CC-RCC patients									
XX	Sequence	909 BP;	218 A;	292 C;	256 G;	143 T;	0 U;	0 Other;		
XX	Query Match	36.5%; Score 552; DB 10; Length 909;								
XX	Best Local Similarity	100.0%; Pred. No. 8.5e-243;								
SQ	Matches	552;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	242	ACTCTGGACATCTCGAGAAAGCTGTAGTCTTGGAATTAAC	TTTGAAACAC	TTTAAAGCT	301					
Db	181	ACTCTGGACATCTCGAGAAAGCTGTAGTCTTGGAATTAAC	TTTGAAACAC	TTTAAAGCT	240					
QY	302	TTAACCGCCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGA	361							
Db	241	TTAACCGCCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGA	300							
QY	362	TCTCTGAAATCGCCCATTCAGTCCGACTTGGATCGCTTCCACTCGGATTTCAAACATGC	421							
Db	301	TCTCTGAAATCGCCCATTCAGTCCGACTTGGATCGCTTCCACTCGGATTTCAAACATGC	360							
QY	422	GCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGG	481							
Db	361	GCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGG	420							
QY	482	TGTGTCCAGTCAACAGGTCCCTCTGAGCAAAAGGCACCGGCTCCCTCGGCCCGGGTCC	541							
Db	421	TGTGTCCAGTCAACAGGTCCCTCTGAGCAAAAGGCACCGGCTCCCTCGGCCCGGGTCC	480							
QY	542	CTGTTGACTCAACAGGTCCCTCTGAGCAAAAGGCACCGGCTCCCTCGGCCCGGGTCC	601							
Db	481	CTGTTGACTCAACAGGTCCCTCTGAGCAAAAGGCACCGGCTCCCTCGGCCCGGGTCC	540							
QY	602	GCGGCGCCCTGCTGGAGCGCGGGSCAGAAGCTGGAGCCCTCGCCTACTGCGTG	661							
Db	541	GCGGCGCCCTGCTGGAGCGCGGGSCAGAAGCTGGAGCCCTCGCCTACTGCGTG	600							
QY	662	CCCGTCATCCAGCGGACTCAGCCCGAGCTCGCCCGCGAGAACGACACGGACACC	721							
Db	601	CCCGTCATCCAGCGGACTCAGCCCGAGCTCGCCCGCGAGAACGACACGGACACC	660							
QY	722	GACAGCGGCTACGGCGGCGAAGCCGAGSCCGGCGGACCCGCGAGAAAGGCGGCGG	781							
Db	661	GACAGCGGCTACGGCGGCGAAGCCGAGSCCGGCGGACCCGCGAGAAAGGCGGCGG	720							
QY	782	GGGGCGAGCCGC	793							
Db	721	GGGGCGAGCCGC	732							
RESULT 10										
AAS26175	ID	AAS26175 standard; cDNA; 498 BP.								
XX	AC	AAS26175;								
XX	DT	07-NOV-2001 (first entry)								
XX	DE	Human cDNA encoding a novel secreted protein, Seq ID 354.								
XX	KW	Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;								
XX	KW	cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;								
XX	KW	antibacterial; virucide; fungicide; ophthalmological; vulnerary;								
XX	KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;								
XX	KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;								
XX	KW	cerebral ischaemia; angiogenesis; nervous system disorder;								
XX	KW	Alzheimer's disease; infection; ocular disorder; corneal infection;								
XX	KW	wound healing; epithelial cell proliferation; skin ageing; food additive;								
XX	OS	preservative; antiproliferative.								
XX	OS	Homo sapiens.								
XX	PN	W020015322-A2.								
XX	XX									

PD	02-AUG-2001.	
XX	17-JAN-2001;	2001WO-US001341.
XX	31-JAN-2000;	2000US-0179065P.
PR	04-FEB-2000;	2000US-0180628P.
PR	24-FEB-2000;	2000US-0184664P.
PR	02-MAR-2000;	2000US-0186350P.
PR	16-MAR-2000;	2000US-0189874P.
PR	17-MAR-2000;	2000US-0190076P.
PR	18-APR-2000;	2000US-0198123P.
PR	19-MAY-2000;	2000US-0205515P.
PR	07-JUN-2000;	2000US-0209467P.
PR	28-JUN-2000;	2000US-0214886P.
PR	30-JUN-2000;	2000US-0215135P.
PR	07-JUL-2000;	2000US-0216647P.
PR	07-JUL-2000;	2000US-0216880P.
PR	11-JUL-2000;	2000US-0217487P.
PR	11-JUL-2000;	2000US-0217496P.
PR	14-JUL-2000;	2000US-0218290P.
PR	26-JUL-2000;	2000US-0220963P.
PR	26-JUL-2000;	2000US-0220964P.
PR	14-AUG-2000;	2000US-0224518P.
PR	14-AUG-2000;	2000US-0224519P.
PR	14-AUG-2000;	2000US-0225213P.
PR	14-AUG-2000;	2000US-0225214P.
PR	14-AUG-2000;	2000US-0225266P.
PR	14-AUG-2000;	2000US-0225267P.
PR	14-AUG-2000;	2000US-0225268P.
PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225757P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	18-AUG-2000;	2000US-0226279P.
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-0226868P.
PR	22-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-0227009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	12-SEP-2000;	2000US-0232081P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.

DR P-PSDB; AAU16188.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
PS Claim 1; SEQ ID NO 354; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins and in diagnostic
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. cardiac
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cerebral
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 31.5%; Score 476; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 6.4e-208;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 134 GATACCTACAAATTACCGCAGATTAAATAGAAAAGAAAGAGACCGAATTAATGAA 193
Db |||||||
QY 194 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGGACAT 253
Db |||||||
QY 254 CTGGAGAAAGCTGTAGCTTTGGAATTAACCTTGAACACTTAAAGCTTTAACCGCCTTA 313
Db |||||||
QY 314 ACCGAGCAACAGCATCAGAGATAATTGCTTTACAGATGGGAGCGATCTCTGAAATCG 373
Db |||||||
QY 374 CCCATTCACTCCGACTTGGATGGTTCCACTCGGGATTCAAACATGCGCCAAAGAGTC 433
Db |||||||
QY 434 TTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGAGCCCGGTGTGTCCAGCTG 493
Db |||||||
QY 494 ATCAACCACTTGACCGCGTGGCCACCCAGTCTCTTGCCACCCCGCAGCTGTTGACTCAA 553
Db |||||||
QY 554 CAGGTCCCTCTGAGCAAGGACCGGGCGTCCCTCGGCGCCCGGGTCCCGCGCGCG 609
Db |||||||
QY 609 CAGGTCCCTCTGAGCAAGGACCGGGCGTCCCTCGGCGCCCGGGTCCCGCGCGCG 696
Db |||||||

RESULT 11
ABX73516
ID ABX73516 standard; DNA; 498 BP.

XX ABX73516;
AC
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polynucleotide #344.
XX
KW Human; Gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
XX US2002132753-A1.
XX
PD 19-SEP-2002.
XX
XX 17-JAN-2001; 2001US-00764864.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-147444/14.
DR P-PSDB; ABU55256.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 354; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 498 BP; 142 A; 137 C; 115 G; 104 T; 0 U; 0 Other;

Query Match 31.5%; Score 476; DB 8; Length 498;
Best Local Similarity 100.0%; Pred. No. 6.4e-208;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAATGAA 193
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAATGAA 80
QY 194 TGCATTGCTCAGCTGAAAGATTTACTGCCTGAACATCTGAAATTGACAACCTCTGGGACAT 253
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
81 TGCATTGCTCAGCTGAAAGATTTACTGCCTGACATCTGAAATTGACAACCTCTGGGACAT 140
QY 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTGAACACACTTAAAGCTTTAACCGCCTTA 313
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
141 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTGAACACACTTAAAGCTTTAACCGCCTTA 200
QY 314 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATGGGGAGCGATCTCTGAAATCG 373
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
201 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATGGGGAGCGATCTCTGAAATCG 260
QY 374 CCCATTCACTCCGACTTGGATGCGTTCCACTCGGGATTTCAAACATGCCCAAAGAGTC 433
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
261 CCCATTCACTCCGACTTGGATGCGTTCCACTCGGGATTTCAAACATGCCCAAAGAGTC 320
QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGGTGTGTCCAGCTG 493
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
321 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGGTGTGTCCAGCTG 380
QY 494 ATCAACCACCTTGACGCGCGTGGCCACCCAGTCTTGCCCAACCCCGAGCTGTGACTCAA 553
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
381 ATCAACCACCTTGACGCGCGTGGCCACCCAGTCTTGCCCAACCCCGAGCTGTGACTCAA 440
QY 554 CAGGTCCCTCTGAGCAAAAGGCACCGGCGTCCCTCGGCCCGCGGTTCGGCGCCGC 609
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
441 CAGGTCCCTCTGAGCAAAAGGCACCGGCGTCCCTCGGCCCGCGGTTCGGCGCCGC 496

AAS26590
ID AAS26590 standard; cDNA; 454 BP.
XX
AC AAS26590;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 769.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.

PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488783/53.
DR P-PSDB; AAU16603.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
PS Claim 1; SEQ ID NO 769; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
Query Match 24.2%; Score 366; DB 4; Length 454;
Best Local Similarity 99.8%; Pred. No. 1.9e-157;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 151 GCACAGATTAAATAGAAAAGAAAAGAGAGACCGAATTATGATGCTCAGCTGAA 210
Db 38 GCACAGATTAAATAGAAAAGAAAAGAGAGACCGAATTATGATGCTCAGCTGAA 97
QY 211 AGATTACTGCCTGAACATCTGAATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGT 270
Db 98 AGATTACTGCCTGAACATCTGAATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGT 157
QY 271 CTTGGAATTAACTTTGAAACACTTAAAGCTTTAACCGCCTTAACCGAGCAACAGCATCA 330
Db 158 CTTGGAATTAACTTTGAAACACTTAAAGCTTTAACCGCCTTAACCGAGCAACAGCATCA 217
QY 331 GAAGATAATTGCTTTACAGAAATGGGAGCGGATCTCTGAAATGCCCATTCAGTCCGACTT 390

Db 218 GAAGATAATTGCTTTACAGAAATGGGAGCGGATCTCTGAAATCGCCCATTCAGTCCGACTT 277
QY 391 GGATGCGTTCCACTCGGGATTTCAAACATGCGCAAAAGAAAGTCTTTGCAATACCTCTCCCG 450
Db 278 GGATGCGTTCCACTCGGGATTTCAAACATGCGCAAAAGAAAGTCTTTGCAATACCTCTCCCG 337
QY 451 GTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTGTCCAGCTGATCAACCACTTGCAACGC 510
Db 338 GTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTGTCCAGCTGATCAACCACTTGCAACGC 397
QY 511 CGTGGCCACCCAGTTCTTGCCCAACCCCGCAGCTGTTGACTCAACAGGTCCTCTGAG 567
Db 398 CGTGGCCACCCAGTTCTTGCCCAACCCCGCAGCTGTTGACTCAACAGGTCCTCTGAG 454

RESULT 13

ABX73931
ID ABX73931 standard; DNA; 454 BP.

XX AC ABX73931;

XX DT 18-MAR-2003 (first entry)

XX DE Human novel polynucleotide #759.

XX KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.

XX OS Homo sapiens.

XX PN US2002132753-A1.

XX PD 19-SEP-2002.

XX PF 17-JAN-2001; 2001US-00764864.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 22-AUG-2000; 2000US-0226868P.

XX PR 30-AUG-2000; 2000US-0228924P.

XX PR 01-SEP-2000; 2000US-0229287P.

XX PR 01-SEP-2000; 2000US-0229343P.

XX PR 01-SEP-2000; 2000US-0229344P.

XX PR 01-SEP-2000; 2000US-0229345P.

XX PR 05-SEP-2000; 2000US-0229509P.

XX PR 05-SEP-2000; 2000US-0229513P.

XX PR 08-SEP-2000; 2000US-0231413P.

XX PR 21-SEP-2000; 2000US-0234223P.

XX PR 21-SEP-2000; 2000US-0234274P.

XX PR 25-SEP-2000; 2000US-0234997P.

XX PR 27-SEP-2000; 2000US-0235834P.


```
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR P-PSDB; ABU55671.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 769; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 454 BP; 137 A; 115 C; 100 G; 100 T; 0 U; 2 Other;
Query Match 24.2%; Score 366; DB 8; Length 454;
Best Local Similarity 99.8%; Pred. No. 1.9e-157;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 151 GCACAGATTAAAGAAAGAAAGAGACCGGAATTATGAATGCATTGCTCAGCTGAA 210
Db |||||||
38 GCACAGATTAAAGAAAGAAAGAGACCGGAATTATGAATGCATTGCTCAGCTGAA 97
QY 211 AGATTTACTGCTGAACATCTGAAATTGACAACCTCTGGGACATCTGGAGAAAGCTGTAGT 270
Db |||||||
98 AGATTTACTGCTGAACATCTGAAATTGACAWCTCTGGGACATCTGGAGAAAGCTGTAGT 157
QY 271 CTTGGAATTAACCTTTGAAACACCTTAAAGCTTTAACCGCTTTAACCGAGCAACAGCATCA 330
Db |||||||
158 CTTGGAATTAACCTTTGAAACACCTTAAAGCTTTAACCGCTTTAACCGAGCAACAGCATCA 217
QY 331 GAAGATAATTGCTTTACAGAATGGGGAGCGGATCTCTGAATTCGCCCATTCAGTCGACTT 390
Db |||||||
218 GAAGATAATTGCTTTACAGAATGGGGAGCGGATCTCTGAATTCGCCCATTCAGTCGACTT 277
```

```
QY 391 GGATGCGTTCCACTCGGATTTCAAACATGCGCCAAAGAAAGTCTTTCGAAATACCTCTCCCG 450
Db |||||||
278 GGATGCGTTCCACTCGGATTTCAAACATGCGCCAAAGAAAGTCTTTCGAAATACCTCTCCCG 337
QY 451 GTTTGAGAGCTGGACACCCAGGGAGCCGCGGTGTGTCCAGCTGATCAACCACTTGACACGC 510
Db |||||||
338 GTTTGAGAGCTGGACACCCAGGGAGCCGCGGTGTGTCCAGCTGATCAACCACTTGACACGC 397
QY 511 CGTGCCACCCAGTTCTTGCCACCCCGCAGCTGTGTGACTCAACAGGTCCCTCTGAG 567
Db |||||||
398 CGTGCCACCCAGTTCTTGCCACCCCGCAGCTGTGTGACTCAACAGGTCCCTCTGAG 454
RESULT 14
ACD19339
ID ACD19339 standard; cDNA; 628 BP.
XX
AC ACD19339;
XX
DT 25-AUG-2003 (first entry)
XX
DE cDNA encoding novel human protein #19.
XX
KW Human; NOV; Gene therapy; endocrine related disease; diabetes;
KW metabolism-related disease; obesity; central nervous system disorder;
KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW stroke; infection; gene; ss.
XX
OS Homo sapiens.
XX
XX WO2003023002-A2.
XX
XX 20-MAR-2003.
XX
XX 09-SEP-2002; 2002WO-US028539.
XX
XX 07-SEP-2001; 2001US-0318120P.
XX 07-SEP-2001; 2001US-0318130P.
XX 10-SEP-2001; 2001US-0318430P.
XX 17-SEP-2001; 2001US-0322636P.
XX 17-SEP-2001; 2001US-0322781P.
XX 17-SEP-2001; 2001US-0322816P.
XX 17-SEP-2001; 2001US-0322817P.
XX 19-SEP-2001; 2001US-0323519P.
XX 20-SEP-2001; 2001US-0323631P.
XX 20-SEP-2001; 2001US-0323636P.
XX 25-SEP-2001; 2001US-0324969P.
XX 25-SEP-2001; 2001US-0325091P.
XX 26-SEP-2001; 2001US-0324990P.
XX 17-APR-2002; 2002US-0373212P.
XX 06-SEP-2002; 2002US-00236177.
XX
(CURA-) CURAGEN CORP.
XX
PI Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
PI Lepley DM, Edinger SR, Burgess CE;
XX
XX WPI; 2003-313242/30.
DR P-PSDB; ABO14646.
XX
PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT and polynucleotides, useful in gene therapy, e.g. for treating or
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
```

PT stroke or infections.
XX
PS Claim 20; Page 125-126; 586pp; English.
XX
CC The invention describes a new isolated polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful as therapeutics,
CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This sequence encodes a novel human NOV protein
XX
SQ Sequence 628 BP; 189 A; 157 C; 147 G; 135 T; 0 U; 0 Other;
Query Match 22.6%; Score 342; DB 10; Length 628;
Best Local Similarity 99.7%; Pred. No. 1.9e-146;
Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAATGAA 193
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
133 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAATGAA 192
QY 194 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGGACAT 253
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
193 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGGACAT 252
QY 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTAACCGCCTTA 313
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
253 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTAACCGCCTTA 312
QY 314 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGGATCTCTGAATCG 373
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
313 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGGATCTCTGAATCG 372
QY 374 CCCATTCACTCCGACTTGGATGCGTCCACTCGGATTTCAACATCGCCCAAGAGTC 433
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
373 CCCATTCACTCCGACTTGGATGCGTCCACTCGGATTTCAACATCGCCCAAGAGTC 432
QY 434 TTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGAGCCGCGTGTGTCCAGCTG 493
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
433 TTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGAGCCGCGTGTGTCCAGCTG 492
QY 494 ATCAACCACTTGCACGCCGCTGGCCACCCAGTTC 526
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
493 ATCAACCACTTGCACGCCGCTGGCCACCCAGTTC 525
RESULT 15
AAS24787
ID AAS24787 standard; cDNA; 285 BP.
XX
AC AAS24787;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human ovarian PCR-subtracted cDNA library clone #968.
XX
KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.

XX Homo sapiens.
OS
XX WO200157207-A2.
PN
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US003733.
PF
XX 04-FEB-2000; 2000US-0180403P.
PR
XX 28-MAR-2000; 2000US-0192745P.
XX
XX (CORI-) CORIXA CORP.
PA
XX Algate PA, Mannion J;
PI
XX WPI; 2001-488879/53.
DR
XX
XX New polynucleotides encoding ovarian tumor proteins, useful for treating
PT ovarian cancer, and as probes, primers, and markers of cancer
PT progression.
XX
PS Example 1; Page 264; 378pp; English.
XX
CC The invention comprises compositions used for the therapy and diagnosis
CC of ovarian cancer. The compositions comprise one or more ovarian tumour
CC proteins, their associated polynucleotides, or immunogenic portions of
CC the proteins. The ovarian tumour polynucleotides and polypeptides are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein. They are also useful for inhibiting the development of cancer in
CC a patient with an ovarian tumour DNA or protein by incubating isolated T-
CC cells allowing them to proliferate, and administering to the patient. The
CC sequences can be used as markers for cancer, for example, to monitor
CC ovarian cancer progression. Probes and primers are useful in nucleic acid
CC hybridisation, in detecting the presence of complementary sequences in a
CC given sample, for preparing mutant species and for preparing other
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549
CC represent human ovarian tumour protein cDNA clones
XX
SQ Sequence 285 BP; 82 A; 68 C; 66 G; 68 T; 0 U; 1 Other;
Query Match 18.1%; Score 274; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 3e-115;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 217 ACTGCCTGAACATCTGAAATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGA 276
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
12 ACTGCCTGAACATCTGAAATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGA 71
QY 277 ATTAACCTTTGAAACACTTAAAGCTTTAACCGCCTTAACCGAGCAACAGCATCAGAAGAT 336
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
72 ATTAACCTTTGAAACACTTAAAGCTTTAACCGCCTTAACCGAGCAACAGCATCAGAAGAT 131
QY 337 AATTGCTTTACAGAAATGGGAGCGGATCTCTGAAATCGCCCATTCAGTCCGACTTGGATGC 396
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
132 AATTGCTTTACAGAAATGGGAGCGGATCTCTGAAATCGCCCATTCAGTCCGACTTGGATGC 191
QY 397 GTTCCACTCGGGATTTCAAACATGCGCCAAAAGAGTCTTGCAATACCTCTCCCGGTTTGA 456
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
192 GTTCCACTCGGGATTTCAAACATGCGCCAAAAGAGTCTTGCAATACCTCTCCCGGTTTGA 251
QY 457 GAGCTGGACACCCAGGAGCCGCGTGTGTCCAG 490
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
252 GAGCTGGACACCCAGGAGCCGCGTGTGTCCAG 285

Search completed: December 17, 2004, 05:53:30
Job time : 693.631 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 17, 2004, 03:15:04 ; Search time 6271.88 Seconds
(without alignments)
11392.870 Million cell updates/sec

Title: US-10-078-650-11
Perfect score: 1511
Sequence: 1 catggagcaaggaattcctc.....gtgagaagttaataaccct 1511

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1511	100.0	1511	6	BD011867
2	1378	91.2	3641	6	BD011858
3	1378	91.2	3641	9	AB044088
4	1327	87.8	3431	9	BC025968
5	1323	87.6	2304	6	AX591648
6	1159	76.7	204228	9	AC022509
7	725	48.0	3274	6	BD176928
8	552	36.5	909	6	CQ727066
9	274	18.1	285	6	AX209128
10	274	18.1	397	6	AX198590
11	143	9.5	3455	4	AY204568
12	128	8.5	260	6	CQ687160
13	128	8.5	260	6	CQ688754
14	128	8.5	260	6	CQ695847
15	128	8.5	260	6	CQ687963
16	128	8.5	265	6	CQ686173
17	128	8.5	278	6	CQ695845
18	128	8.5	502	6	AX892348
19	128	8.5	502	6	BD027881

20	113	7.5	5245	4	AY204567	AY204567	Canis fam
21	77	5.1	311	6	CQ696097	CQ696097	Sequence
22	54	3.6	692	10	BC037742	BC037742	Mus muscu
23	54	3.6	1421	6	BD011868	BD011868	Novel bHL
24	54	3.6	1421	10	AB044090	AB044090	Mus muscu
25	54	3.6	3101	10	AF009329	AF009329	Rattus no
26	46	3.0	300	6	BD220060	BD220060	Human gen
27	46	3.0	13281	10	AB126167	AB126167	Mus muscu
28	46	3.0	212716	2	AC144936	AC144936	Mus muscu
29	46	3.0	272078	2	AC138117	AC138117	Mus muscu
30	44	2.9	236385	2	AC110349	AC110349	Rattus no
31	33	2.2	439	10	MMBMPG1	AF100904	Mus muscu
32	33	2.2	439	10	MMGDF11501	AF028335	Mus muscu
33	33	2.2	664	10	MMU131846	AJ131846	Mus muscu
34	33	2.2	671	4	AY677108	AY677108	Dasyopus s
35	33	2.2	161757	10	AC128700	AC128700	Mus muscu
36	33	2.2	204856	2	AC119275	AC119275	Mus muscu
37	33	2.2	211163	2	AC141508	AC141508	Rattus no
38	32	2.1	160	9	S70397	S70397	{FRA16A, fo
39	32	2.1	390	4	AF083096	AF083096	Canis fam
40	32	2.1	393	4	AF083095	AF083095	Felis cat
41	32	2.1	3170	8	AK066310	AK066310	Oryza sat
42	32	2.1	8147	6	AR203334	AR203334	Sequence
43	32	2.1	8147	6	AX777237	AX777237	Sequence
44	32	2.1	8147	9	HSU85962	U85962	Homo sapien
45	32	2.1	35687	9	AC004779	AC004779	Homo sapi

ALIGNMENTS

RESULT 1
BD011867
LOCUS BD011867 1511 bp DNA linear PAT 02-AUG-2002
DEFINITION Novel bHLH type transcriptional gene, DEC2.
ACCESSION BD011867
VERSION BD011867.1 GI:22092056
KEYWORDS WO 0114551-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1511)
AUTHORS Fujimoto,K., Shin,M. and Kato,Y.
TITLE Novel bHLH type transcriptional gene, DEC2
JOURNAL Patent: WO 0114551-A 10 01-MAR-2001;
CHUGAI PHARM CO LTD,KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
COMMENT OS Homo sapiens (human)
PN WO 0114551-A/10
PD 01-MAR-2001
PF 19-JUN-2000 WO 2000JP003991
PR 19-AUG-1999 JP 99P 233286
PI KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
PC C12N15/12,C07L14/475,C07K16/18
CC
FH
FT

FEATURES
source Location/Qualifiers
1..1511
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1511; DB 6; Length 1511;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGACGAGGAATTCCTCATTTCGAAGAGACAGTCTACTGGAACATAGATTTTAT 60

Db 1 CATGGACGAGGAATTCCTCATTTCGAAGAGACAGTCTACTGGAACATAGATTTTAT 60

QY 61 AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAACGAGACGA 120

Db 61 AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGA 120
Qy 121 CACCAAGGTAAGTGATACCTACAAATTACCGGCACAGATTAAATAGAAAAGAAAAGAGAGA 180
Db 121 CACCAAGGTAAGTGATACCTACAAATTACCGGCACAGATTAAATAGAAAAGAAAAGAGAGA 180
Qy 181 CCGAATTAAATGAATGCTCAGCTGAAAGATTACTGCTGAAACATCTGAAATTGAC 240
Db 181 CCGAATTAAATGAATGCTCAGCTGAAAGATTACTGCTGAAACATCTGAAATTGAC 240
Qy 241 AACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACTTTGAAACACTTAAAGC 300
Db 241 AACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACTTTGAAACACTTAAAGC 300
Qy 301 TTTAAACCGCCTTAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAAATGGGAGCG 360
Db 301 TTTAAACCGCCTTAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAAATGGGAGCG 360
Qy 361 ATCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTCCACTCGGGATTTCAAACATG 420
Db 361 ATCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTCCACTCGGGATTTCAAACATG 420
Qy 421 CGCCAAAGAAAGTCTTGCATACCTCTCCCGTTTGAGAGCTGGACACCCAGGGAGCCGCG 480
Db 421 CGCCAAAGAAAGTCTTGCATACCTCTCCCGTTTGAGAGCTGGACACCCAGGGAGCCGCG 480
Qy 481 GTGTCTCAGCTGATCAACCACTTGACGCGCTGGCCACCCAGTTCTTGCCACCCCGCA 540
Db 481 GTGTCTCAGCTGATCAACCACTTGACGCGCTGGCCACCCAGTTCTTGCCACCCCGCA 540
Qy 541 GCTGTGACTCAACAGGTCCTCTGAGCAAGGCAACGGCGCTCCCTCGGCGCGCGGTC 600
Db 541 GCTGTGACTCAACAGGTCCTCTGAGCAAGGCAACGGCGCTCCCTCGGCGCGCGGTC 600
Qy 601 CGCGCGCGCCCTGCTGAGCGCGCGGGGAGAGCTGGAGCCCTCGCTACTGCGT 660
Db 601 CGCGCGCGCCCTGCTGAGCGCGCGGGGAGAGCTGGAGCCCTCGCTACTGCGT 660
Qy 661 GCCCGTCATCCAGCGGACTCAGCCAGCGCGGAGCTCGCCGCGAGAACGACGACGAC 720
Db 661 GCCCGTCATCCAGCGGACTCAGCCAGCGCGGAGCTCGCCGCGAGAACGACGACGAC 720
Qy 721 CGACAGCGGCTACGGCGGCGAAGCCGAGGCGCGGCGGAGCGCGGAGAAAGGCGC 780
Db 721 CGACAGCGGCTACGGCGGCGAAGCCGAGGCGCGGCGGAGCGCGGAGAAAGGCGC 780
Qy 781 GGGGGCGAGCGCGCTACCAATCAAGCAGGAGGCTCCCGGGGAGGACTCGCGCGGCCCAA 840
Db 781 GGGGGCGAGCGCGCTACCAATCAAGCAGGAGGCTCCCGGGGAGGACTCGCGCGGCCCAA 840
Qy 841 GAGGATGAAGCTGGATTCCCGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCG 900
Db 841 GAGGATGAAGCTGGATTCCCGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCG 900
Qy 901 GCGCGCGCGAGCGCGCTTCTGGGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 960
Db 901 GCGCGCGCGAGCGCGCTTCTGGGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 960
Qy 961 CGACGCGCGCCTGCTCAGCTCGTGGTGGGCTTGGCGGAGCGGAGCGCGCGCTTCCC 1020
Db 961 CGACGCGCGCCTGCTCAGCTCGTGGTGGGCTTGGCGGAGCGGAGCGCGCGCTTCCC 1020
Qy 1021 GCAGCGCGCGCGCGCGCTTCTGCGCTGCGCTTCTGCTTCTGCTTCTGCTTCTGCT 1080
Db 1021 GCAGCGCGCGCGCGCGCTTCTGCGCTGCGCTTCTGCTTCTGCTTCTGCTTCTGCT 1080
Qy 1081 AGCTGCGCGCTACGTGACGCGCTTCTGGACAAGAGCGCGCTGGAGAAATGTGTACCC 1140
Db 1081 AGCTGCGCGCTACGTGACGCGCTTCTGGACAAGAGCGCGCTGGAGAAATGTGTACCC 1140
Qy 1141 GCGCGCGCTGCGCGCGCTTCCGCTGCTATACCCCGCGCATCCCGCGCGCGCGCGCAGC 1200

Db 1141 GCGCGCGCTGCGCGCGCTTCCGCTGCTATATACCCCGCATCCCGCGCGCGCGCAGC 1200
Qy 1201 CCGCGCAGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCTTCCCGCTGCTTCTCGGT 1260
Db 1201 CCGCGCAGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCTTCCCGCTGCTTCTCGGT 1260
Qy 1261 GTTGTGCGCGCTTCCGAGAAAGGGGGCGCGCGCGCGCGCGCGCTTCCCTGCGCAGGAGT 1320
Db 1261 GTTGTGCGCGCTTCCGAGAAAGGGGGCGCGCGCGCGCGCGCGCTTCCCTGCGCAGGAGT 1320
Qy 1321 GCGCGCGCTTGGGGCGCGCACCCCGAGCACCCCGACGCGCGCACCCACTGCGCTTCGC 1380
Db 1321 GCGCGCGCTTGGGGCGCGCACCCCGAGCACCCCGACGCGCGCACCCACTGCGCTTCGC 1380
Qy 1381 CCGCGCGCGCGAGCGCGGAAACCCGAGAGCTTGTCTCAGGAAGATCCCTCGCAGCCAGG 1440
Db 1381 CCGCGCGCGCGAGCGCGGAAACCCGAGAGCTTGTCTCAGGAAGATCCCTCGCAGCCAGG 1440
Qy 1441 AAAGGAAGCTCCCTGAATCCTTGGTCCCGTCCCGAAGCGGAGGTTCAAGCAGAGTGAGAAGT 1500
Db 1441 AAAGGAAGCTCCCTGAATCCTTGGTCCCGTCCCGAAGCGGAGGTTCAAGCAGAGTGAGAAGT 1500
Qy 1501 TAAATATACCT 1511
Db 1501 TAAATATACCT 1511

RESULT 2
BD011858
LOCUS BD011858 3641 bp DNA linear PAT 02-AUG-2002
DEFINITION Novel bHLH type transcriptional gene, DEC2.
ACCESSION BD011858
VERSION BD011858.1 GI:22092047
KEYWORDS WO 0114551-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3641)
AUTHORS Fujimoto,K., Shin,M. and Kato,Y.
TITLE Novel bHLH type transcriptional gene, DEC2
JOURNAL Patent: WO 0114551-A 1 01-MAR-2001;
CHUGAI PHARM CO LTD,KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
COMMENT OS Homo sapiens (human)
PN WO 0114551-A/1
PD 01-MAR-2001
PF 19-JUN-2000 WO 2000JP003991
PR 19-AUG-1999 JP 99P 233286
PI KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
PC C12N15/12,C07L14/475,C07K16/18
CC

FEATURES
source
FH Key Location/Qualifiers
FT CDS (135)..(1580).
1..3641
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 91.2%; Score 1378; DB 6; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAAGAGACCGGAATTAATGAA 193
Db 261 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAAGAGACCGGAATTAATGAA 320
Qy 194 TGCATTGCTCAGCTGAAAGATTACTGCTGCAACATCTGAAATTTGACAACTCTGGGACAT 253
Db 321 TGCATTGCTCAGCTGAAAGATTACTGCTGCAACATCTGAAATTTGACAACTCTGGGACAT 380
Qy 254 CTGGAGAAAGCTGTAGTCTTGAATTAACCTTTGAACACTTAAAGCTTTAACCCTTA 313

Db 381 CTGGAGAAAGCTGTAGTCTTTGGAATTAACCTTTGAAACACTTAAAGCTTTAACCGCCTTA 440
QY 314 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATGGGAGCGATCTCTGAAATCG 373
Db 441 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATGGGAGCGATCTCTGAAATCG 500
QY 374 CCCATTCACTCCGACTTGGATGCGTTCCACTCGGGATTTCACATCGCCCAAGAGTC 433
Db 501 CCCATTCACTCCGACTTGGATGCGTTCCACTCGGGATTTCACATCGCCCAAGAGTC 560
QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGACACCCAGGGAGCCCGGTTGTGTCCAGCTG 493
Db 561 TTGCAATACCTCTCCCGGTTTGAGAGCTGACACCCAGGGAGCCCGGTTGTGTCCAGCTG 620
QY 494 ATCAACCACTTGACACCGCGTGGCCACCCAGTTCTTGCCCAACCCCGAGCTGTTGACTCA 553
Db 621 ATCAACCACTTGACACCGCGTGGCCACCCAGTTCTTGCCCAACCCCGAGCTGTTGACTCA 680
QY 554 CAGGTCCCTCTGAGCAAGGACCCGCGCTCCCTCGGCCCGCGGTCGCGCGCGCCGCC 613
Db 681 CAGGTCCCTCTGAGCAAGGACCCGCGCTCCCTCGGCCCGCGGTCGCGCGCGCCGCC 740
QY 614 TGCCTGGAGCGCGCGGCGAGAGCTGAGACCCCTCGCTTCTGCTGCTGCTGCTGCTGCTG 673
Db 741 TGCCTGGAGCGCGCGGCGAGAGCTGAGACCCCTCGCTTCTGCTGCTGCTGCTGCTGCTG 800
QY 674 CGGACTCAGCCAGCGCGGAGCTCGCGCGGAGAACGACACCGACACCGACAGCGGCTAC 733
Db 801 CGGACTCAGCCAGCGCGGAGCTCGCGCGGAGAACGACACCGACACCGACAGCGGCTAC 860
QY 734 GCGGGGAGCGCGGAGCGCGGAGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCG 793
Db 861 GCGGGGAGCGCGGAGCGCGGAGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCG 920
QY 794 GTCACCATCAAGCAGGAGCTCCCGGGGAGGACTCGCGGCGCGGCGCGGCGCGGCGCGG 853
Db 921 GTCACCATCAAGCAGGAGCTCCCGGGGAGGACTCGCGGCGCGGCGCGGCGCGGCGCGG 980
QY 854 GATTCCCGCGCGCGGAGCGCGGAGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGG 913
Db 981 GATTCCCGCGCGCGGAGCGCGGAGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGG 1040
QY 914 GCGCTTCTGCGGCGCGGAGCGCGGAGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGG 973
Db 1041 GCGCTTCTGCGGCGCGGAGCGCGGAGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGG 1100
QY 974 CTCAGTCGCTGGTGGCGTTTCGGCGGAGCGCGGAGCGCGGCGCGGCGCGGCGCGGCGG 1033
Db 1101 CTCAGTCGCTGGTGGCGTTTCGGCGGAGCGCGGAGCGCGGCGCGGCGCGGCGCGGCGG 1160
QY 1034 GCGCGGCGCGGCTTCTGCTGCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 1093
Db 1161 GCGCGGCGCGGCTTCTGCTGCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 1220
QY 1094 GTGCAGCGCTTCTGGAACAAGCGGCGCTGGAGAGTATCTGTACCCGCGGCGGCGGCTGCC 1153
Db 1221 GTGCAGCGCTTCTGGAACAAGCGGCGCTGGAGAGTATCTGTACCCGCGGCGGCGGCTGCC 1280
QY 1154 GCGCGGCTTCTGCTATACCCGCGGCTATACCCGCGGCTATACCCGCGGCTATACCCGCGG 1213
Db 1281 GCGCGGCTTCTGCTATACCCGCGGCTATACCCGCGGCTATACCCGCGGCTATACCCGCGG 1340
QY 1214 GCGCGGCTTCTGCTGCTGCTGCTTCTGCTGCTGCTTCTGCTGCTGCTTCTGCTGCTGCT 1273
Db 1341 GCGCGGCTTCTGCTGCTGCTGCTTCTGCTGCTGCTTCTGCTGCTGCTTCTGCTGCTGCT 1400
QY 1274 CCGGAGAGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1333
Db 1401 CCGGAGAGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1460
QY 1334 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1393

Db 1461 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1520
QY 1394 CCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGAGCTCCC 1453
Db 1521 CCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGAGCTCCC 1580
QY 1454 TGAATCCTTGCCTCCGAAAGGACGAGGTTCAAGCAGAGTGAGAGTTAAATACCCCT 1511
Db 1581 TGAATCCTTGCCTCCGAAAGGACGAGGTTCAAGCAGAGTGAGAGTTAAATACCCCT 1638
RESULT 3
AB044088
LOCUS Homo sapiens mRNA for bHLH protein DEC2, complete cds. PRI 27-JAN-2001
DEFINITION
ACCESSION AB044088
VERSION AB044088.1 GI:12583668
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
AUTHORS Fujimoto,K., Shen,M., Noshiro,M., Matsubara,K., Shingu,S.,
Honda,K., Yoshida,E., Suardita,K., Matsuda,Y. and Kato,Y.
TITLE Molecular cloning and characterization of DEC2, a new member of
basic helix-loop-helix proteins
JOURNAL Biochem. Biophys. Res. Commun. 280 (1), 164-171 (2001)
MEDLINE 21092582
PUBMED 11162494
REFERENCE 2 (bases 1 to 3641)
AUTHORS Fujimoto,K.
TITLE Direct Submission
JOURNAL Submitted (31-MAY-2000) Katsumi Fujimoto, hiroshima university,
department of biochemistry, school of dentistry; 1-2-3, kasumi,
minami-ku, hiroshima 734-8553, Japan
(E-mail:kfujimo@hiroshima-u.ac.jp, Tel:81822575629,
Fax:81822575629)
FEATURES
Location/Qualifiers
1..3641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1..3641
/gene="DEC2"
CDS 135..1583
/gene="DEC2"
/codon_start=1
/product="bHLH protein DEC2"
/protein_id="BAB21502.1"
/db_xref="GI:12583669"
/translation="MDEGPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDDTKDT
YKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELT.LKHLKALTAL
TEQQHQKIIALQNGERSLSPISQDLDLAFHSGFTCAKEVLQYLSRFESWTPEPRCV
QLINHLHAVATQFLPTPOLLTQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCV
PVIQRTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPFGEDSPA
PKMKLDSRGGSGGGGAGAAALPDPAALALRPDAALSLVAFGGGG
APFPQAAAAAPFCLPFCFLSPSAAAYVQPFLLSGLEKLYPAAAAAPFLLYPGI
PAPAAAAAALAAAAAAPPCLSSVLSPPPEKAGAAAATLLPHEVAPLGAHPQHPH
GRTHLPFAGPREPGNPSSAQEDPSQPGKEAP"
ORIGIN
Query Match 91.2%; Score 1378; DB 9; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAAGAGAGACCGAATTAATGAA 193
Db 261 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAAGAGAGACCGAATTAATGAA 320
QY 194 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATGACAACTCTGGGACAT 253
Db 321 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATGACAACTCTGGGACAT 380

QY	254	CTGAGAAAGCTGTAGTCTTGAATTAACCTTTGAAACACTTAAAGCTTTAAACCGCCTTA	313
Db	381	CTGAGAAAGCTGTAGTCTTGAATTAACCTTTGAAACACTTAAAGCTTTAAACCGCCTTA	440
QY	314	ACCGAGCAACAGCATCAGAAGATAATTCCTTTACAGAAATGGGAGCGATCTCTGAAATCG	373
Db	441	ACCGAGCAACAGCATCAGAAGATAATTCCTTTACAGAAATGGGAGCGATCTCTGAAATCG	500
QY	374	CCCATTCAGTCCGACTTGGATCGCTTCCACTCGGATTTCAAACATGCGCAAAAGAGTC	433
Db	501	CCCATTCAGTCCGACTTGGATCGCTTCCACTCGGATTTCAAACATGCGCAAAAGAGTC	560
QY	434	TTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGGAGCCGCGTGTGTCCAGCTG	493
Db	561	TTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGGAGCCGCGTGTGTCCAGCTG	620
QY	494	ATCAACCACTTGACCGCGTGGCCACCCAGTTCTTGCCACCCCGCAGCTGTGACTCAA	553
Db	621	ATCAACCACTTGACCGCGTGGCCACCCAGTTCTTGCCACCCCGCAGCTGTGACTCAA	680
QY	554	CAGTCCCTCTGAGCAAGGACCGGCGCTCCCTCGGCCCGCGGTCCCGGGCGGCCCCC	613
Db	681	CAGTCCCTCTGAGCAAGGACCGGCGCTCCCTCGGCCCGCGGTCCCGGGCGGCCCCC	740
QY	614	TGCTGGAGCGCGGGGCGAAGCTGGAGCCCTCGCTACTGCGTGCCCGTCATCCAG	673
Db	741	TGCTGGAGCGCGGGGCGAAGCTGGAGCCCTCGCTACTGCGTGCCCGTCATCCAG	800
QY	674	CGGACTCAGCCAGCGCCGAGTCCGCGGAGACACACGACACCGACACCGACCGGCTAC	733
Db	801	CGGACTCAGCCAGCGCCGAGTCCGCGGAGACACACGACACCGACACCGACCGGCTAC	860
QY	734	GGCGGGAAGCCGAGGCCCGCGGACCGCGAGAAAGCAAGCGCGGGGCGAGCCGC	793
Db	861	GGCGGGAAGCCGAGGCCCGCGGACCGCGAGAAAGCAAGCGCGGGGCGAGCCGC	920
QY	794	GTCACCATCAAGCAGGAGCCTCCCGGGGAGGACTCGCGGCGCCCAAGAGGATGAAGCTG	853
Db	921	GTCACCATCAAGCAGGAGCCTCCCGGGGAGGACTCGCGGCGCCCAAGAGGATGAAGCTG	980
QY	854	GATTCGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC	913
Db	981	GATTCGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC	1040
QY	914	GCGCTTCTGGGGCCCGACCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC	973
Db	1041	GCGCTTCTGGGGCCCGACCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC	1100
QY	974	CTCAGCTCGCTGGTGGCGTTTCGGGAGGCGGAGCGCGCGCGCGCGCGCGCGCGCG	1033
Db	1101	CTCAGCTCGCTGGTGGCGTTTCGGGAGGCGGAGCGCGCGCGCGCGCGCGCGCGCG	1160
QY	1034	GCCGCGGCGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1093
Db	1161	GCCGCGGCGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1220
QY	1094	GTGACGCTTCTCTGGACAAGAGCGGCTTGAGAGTATCTGTACCCGCGCGCGCTGCC	1153
Db	1221	GTGACGCTTCTCTGGACAAGAGCGGCTTGAGAGTATCTGTACCCGCGCGCGCTGCC	1280
QY	1154	GCCCCGTTCCGCTGCTATACCCCGCATCCCCCGCGCGCGCGCGCGCGCGCGCGCC	1213
Db	1281	GCCCCGTTCCGCTGCTATACCCCGCATCCCCCGCGCGCGCGCGCGCGCGCGCGCC	1340
QY	1214	GCCGCGCTGCCCT	1273
Db	1341	GCCGCGCTGCCCT	1400
QY	1274	CCCGAGAGGCGGG	1333
Db	1401	CCCGAGAGGCGGG	1460

QY	1334	GCGCCGCAACCCCGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGAGCTCCC	1393
Db	1461	GCGCCGCAACCCCGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGAGCTCCC	1520
QY	1394	CCGGGAACCCCGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGAGCTCCC	1453
Db	1521	CCGGGAACCCCGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGAGCTCCC	1580
QY	1454	TGAATCCTTGGTCCCGAAGGACGGAGGTTCAAGCAGAGTGAGAAGTTAAATACCCCT	1511
Db	1581	TGAATCCTTGGTCCCGAAGGACGGAGGTTCAAGCAGAGTGAGAAGTTAAATACCCCT	1638
RESULT 4			
LOCUS	BC025968	3431 bp	linear
DEFINITION	Homo sapiens basic helix-loop-helix domain containing, class B, 3, mRNA (cDNA clone MGC:39365 IMAGE:4650325), complete cds.		
ACCESSION	BC025968		
VERSION	BC025968.1	GI:19684063	
KEYWORDS	MGC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 3431)		
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Schuler, G.D., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schaefer, C.F., Bhat, N.K., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 3431)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu, Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.		

Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M. P., Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K. T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shooshtari, N., Sisson, I., Sodergren, E., Sonaiker, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Kucherlapati, R. and Gibbs, R.

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 204228)
Worley, K.C.
Direct Submission
Submitted (05-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 204228)
Worley, K.C.
Direct Submission
Submitted (30-NOV-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 30, 2000 this sequence version replaced gi:10945701.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found

at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Option Info -----
Version: 1.01
\$Id: bcm-qual-stats.cc,v 1.6 1998/10/27 00:34:03 kdurbin Exp
kdurbin \$

Options Specified: xfo
XGapFile: ./xgap/hadw.final
OutFile :
FA File : ./xgap/HADW.fa
Sequence start: 10
Sequence end: 204237
Optimistic: 0
Use no qual cols: 0

----- Summary Statistics -----

Contig length: 204228
Phrap values in estimate: 203313
Average error rate (BCM-Phrap estimate): 0.000109636
Fraction of Phrap values less than 40 : 0.0179674
Number of consensus changing edits: 37
Number of N's in consensus : 11
Clip start: 204238
Clip end: 204238
Num reads missing quality: 174
Num reads invalid quality: 0

----- Low Quality Bases -----

Quality	Position	Surrounding Sequence
4	6263	ttttttttt(t)gagacagggg
7	89744	aattaataac(a)agggcagctg
7	89830	ctggtagttt(g)agaccagcct
7	156088	gacaatcaga(a)atgaacttaa
8	6188	aataactaag(c)taaagaaaaa
8	7146	ccactatcac(a)agaatagcat
8	89632	gtcaggaaaa(t)gccatctgtt
8	89633	tcaggaaaaat(g)ccatctgtta
8	89634	caggaaaaatg(c)catctgttac
8	89745	attaataaca(a)gggcagctgg
8	89791	cagcattttt(g)gaggtgagg
8	89792	agcatttttg(g)aggtgaggtt
8	89795	attttgggag(g)ctgaggtggg
8	89829	tctgtagttt(t)gagaccagcc
8	140166	gggggagggg(g)gaggggggag
8	140167	gggggaggggg(g)aggggggag
9	6160	aaaggaaaaa(t)aaatcttggg
9	7030	gtggaagggtg(a)aggggaagca
9	7097	gaggagcgc(a)aggggagaag
9	7099	gggagcgcaa(a)gggagaagt
9	7106	caaaaggaga(a)gtgccacatc
9	7128	gattcatga(c)aactccccca
9	7130	tctcatgaca(a)ctcccccat
9	7135	tgacaactcc(c)ccactatcac
9	7136	gacaactccc(c)caatacacaa
9	7137	acaactcccc(c)actatcacaa
9	7138	caactcccc(c)ctatcacaa
9	7139	aactccccca(c)tatcacaga
9	7144	ccccctatc(a)caagaatagc
9	7148	actatcacaa(g)aatagcatgg
9	7149	ctatcacaa(g)aatagcatgg
9	7150	tatcacaa(g)aatagcatgg
9	89628	ccagtcagg(a)aaatgccatc
9	89768	atggtggctt(a)cagctgtaat
9	89769	tggtggctta(c)agctgtaatc

9	89793	gcatTTTTGGG(a)ggctgaggtg
9	89828	gtctgtagt(t)tgagaccagc
9	89865	aacccatct(c)tacaaaaat
10	6159	gaaaggaaaa(a)taaatcttg
10	6179	ggccccaaa(a)tactaagct
10	6180	ggccccaaa(t)cactaagcta
10	6209	aagtgaagct(g)ggaactgctt
10	6222	aactgcttg(g)gcaaaactgcc
10	6234	caaaactgct(c)ctatttttt
10	6540	agccaccgtg(c)ctaccgctt
10	6542	ccaccgtgc(c)taccgcttc
10	7029	ggtggaaggt(g)aaggggaagc
10	7057	gtctttacaa(g)gtggcaggag
10	7058	tctttacaag(g)tggcaggagg
10	7085	gagagagaga(c)agagggagcg
10	7086	agagagagac(a)gaggagcgc
10	7087	gagagagaca(g)agggagcgca
10	7096	agaggagcg(c)aaaggagaaa
10	7100	ggagcgcaaa(g)ggagaagtgc
10	7101	gagcgcaaa(g)gagaagtgc
10	7102	agcgcaaa(g)agaagtgcga
10	7103	gcgcaaaagg(a)gaagtgcac
10	7104	cgcaaaagg(a)gaagtgcaca
10	7105	gcaaaaggag(a)agtgccacat
10	7131	ctcatgacaa(c)tccccacta
10	7142	tccccacta(t)cacaaagata
Query Match 76.7%; Score 1159; DB 9; Length 204228;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	353	GGGGAGCGATCTCTGAAATCGCCATTTCAGTCCGATTCGATGCGTTCACCTCGGATT 412
Db	44987	GGGGAGCGATCTCTGAAATCGCCATTTCAGTCCGATTCGATGCGTTCACCTCGGATT 44928
QY	413	CAAAATGCGCCAAAGAGTCTTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGG 472
Db	44927	CAAAATGCGCCAAAGAGTCTTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGG 44868
QY	473	GAGCCGCGTGTGTCCAGCTGATCAACCACTTGACGCGCGTGGCCACCCAGTTCTGCCC 532
Db	44867	GAGCCGCGTGTGTCCAGCTGATCAACCACTTGACGCGCGTGGCCACCCAGTTCTGCCC 44808
QY	533	ACCCGCGAGTGTGACTCAACAGTCCCTCTGAGCAAGGCACCGCGCTCCCTCGGCC 592
Db	44807	ACCCGCGAGTGTGACTCAACAGTCCCTCTGAGCAAGGCACCGCGCTCCCTCGGCC 44748
QY	593	GCCGGTCCGCGCGCCCTGCTGGAGCGCGCGGCGAGAGCTGGAGCCCTCGCC 652
Db	44747	GCCGGTCCGCGCGCCCTGCTGGAGCGCGCGGCGAGAGCTGGAGCCCTCGCC 44688
QY	653	TACTGCTGCGCGTTCATCCAGCGGACTCAGCCAGCGCGCGAGCTCGCCGCGAGAACGAC 712
Db	44687	TACTGCTGCGCGTTCATCCAGCGGACTCAGCCAGCGCGCGAGCTCGCCGCGAGAACGAC 44628
QY	713	ACGGACACCGACAGCGGCTACGCGCGGAAGCCGAGCCCGCGGACCCGCGAGAAAGGC 772
Db	44627	ACGGACACCGACAGCGGCTACGCGCGGAAGCCGAGCCCGCGGACCCGCGAGAAAGGC 44568
QY	773	AAAGGCGGGGGCGAGCCCGCTCACCATCAAGCAGGAGCCTCCCGGGGAGGACTCGCCG 832
Db	44567	AAAGGCGGGGGCGAGCCCGCTCACCATCAAGCAGGAGCCTCCCGGGGAGGACTCGCCG 44508
QY	833	GCGCCCAAGAGGATGAAGCTGATTCCTCGCGCGCGGCGAGCGCGCGCCCGGGGGGC 892
Db	44507	GCGCCCAAGAGGATGAAGCTGATTCCTCGCGCGCGGCGAGCGCGCGCCCGGGGGGC 44448
QY	893	GCGCGCGCGCGCGGCGAGCGCGTCTGCGGCGCGACCTGCGCGCGCGCGCGCTG 952
Db	44447	GCGCGCGCGCGCGGCGAGCGCGTCTGCGGCGCGACCTGCGCGCGCGCGCGCTG 44388
QY	953	CTGAGACCCGACGCGCCCTGCTCAGTCTGCTGGTGGCTTGGCGGAGCGGCGCGG 1012

Db	44387	CTGAGACCCGACGCGCCCTGCTCAGCTCGTGGTGGGTTGCGGAGGCGGAGCGCG 44328
QY	1013	CCCTTCCCGCAGCCCGCGCGCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1072
Db	44327	CCCTTCCCGCAGCCCGCGCGCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTG 44268
QY	1073	CCCTTCTCAGCTGCGCGCTACCTGCTGAGCCCTTCTCTGACAAAGAGCGGCTGGAGAGTAT 1132
Db	44267	CCCTTCTCAGCTGCGCGCTACCTGCTGAGCCCTTCTCTGACAAAGAGCGGCTGGAGAGTAT 44208
QY	1133	CTGTACCCGCGCGCGCTGCGCGCGCGCTTCCCGCTGCTATACCCCGGCATCCCGCGCCG 1192
Db	44207	CTGTACCCGCGCGCGCTGCGCGCGCGCTTCCCGCTGCTATACCCCGGCATCCCGCGCCG 44148
QY	1193	GCGGACCCGCGCGAGCCCGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCGCGCTTCCCGCTG 1252
Db	44147	GCGGACCCGCGCGAGCCCGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCGCGCTTCCCGCTG 44088
QY	1253	TCCTCGTGTGTCGCGCGCGCTTCCCGAGAAAGCGGCGCGCGCGCGCGCGCTTCCCGCTG 1312
Db	44087	TCCTCGTGTGTCGCGCGCGCTTCCCGAGAAAGCGGCGCGCGCGCGCGCTTCCCGCTG 44028
QY	1313	CACGAGTGGCGCGCTTGGGCGCGCGACCCCGAGACCCCGAGACCCCGCGCGCGCGCTG 1372
Db	44027	CACGAGTGGCGCGCTTGGGCGCGCGACCCCGAGACCCCGAGACCCCGCGCGCGCGCTG 43968
QY	1373	CCCTTCCCGCGCGCGCGAGCGCGGGGAAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCG 1432
Db	43967	CCCTTCCCGCGCGCGCGAGCGCGGGGAAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCG 43908
QY	1433	CAGCCAGGAAAGAGCTCCCTGAATCCTTGGTCCCGTCCCGAGCGCGCGCGCGCGCTTCCCGCTG 1492
Db	43907	CAGCCAGGAAAGAGCTCCCTGAATCCTTGGTCCCGTCCCGAGCGCGCGCGCGCTTCCCGCTG 43848
QY	1493	TGAGAGTTAAATACCT 1511
Db	43847	TGAGAGTTAAATACCT 43829

RESULT 7	BD176928	3274 bp	DNA	linear	PAT 16-APR-2003
LOCUS	BD176928	Novel bHLH type transcriptional gene, DEC2.			
DEFINITION	BD176928	Novel bHLH type transcriptional gene, DEC2.			
ACCESSION	BD176928.1	GI:30014187			
VERSION	BD176928.1	GI:30014187			
KEYWORDS	JP 2002300878-A/1.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Fujimoto, K., Shin, M. and Kato, Y.				
TITLE	Novel bHLH type transcriptional gene, DEC2				
JOURNAL	Patent: JP 2002300878-A 1 15-OCT-2002; CHUGAI PHARMACEUT CO LTD				
COMMENT	OS Homo sapiens (human) PN JP 2002300878-A/1 PD 15-OCT-2002 PF 19-AUG-1999 JP 1999233286 PI KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10,PC				
FEATURES	source				
	1..3274				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				

FEATURES
source
CORIXA CORPORATION (US)
Location/Qualifiers
1. .285
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 18.1%; Score 274; DB 6; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.9e-130;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 217 ACTGCCTGAACATCTGAAATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGA 276
Db 12 ACTGCCTGAACATCTGAAATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGA 71
QY 277 ATTAACCTTTGAAACACCTTAAAGCTTTAAACCGCTTTAACCGGACCAACAGCATCAGAAGAT 336
Db 72 ATTAACCTTTGAAACACCTTAAAGCTTTAAACCGCTTTAACCGGACCAACAGCATCAGAAGAT 131
QY 337 AATTGCTTTACAGAAATGGGGAGGAGATCTCTGAAATCGGCCATTTCAGTCCGACTTGGATGC 396
Db 132 AATTGCTTTACAGAAATGGGGAGGAGATCTCTGAAATCGGCCATTTCAGTCCGACTTGGATGC 191
QY 397 GTTCCACTCGGGATTTCAAACATCGGCCAAAGAGTCTTGCAATACCTCTCCCGGTTTGA 456
Db 192 GTTCCACTCGGGATTTCAAACATCGGCCAAAGAGTCTTGCAATACCTCTCCCGGTTTGA 251
QY 457 GAGCTGGACACCCAGGAGCGCGGTGTGTCCAG 490
Db 252 GAGCTGGACACCCAGGAGCGCGGTGTGTCCAG 285

RESULT 10
AX198590
LOCUS AX198590 397 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 1045 from Patent WO0151513.
ACCESSION AX198590
VERSION AX198590.1 GI:15388911
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Algate, P.A.
TITLE Ovarian tumor-associated sequences
JOURNAL Patent: WO 0151513-A 1045 19-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1. .397
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 18.1%; Score 274; DB 6; Length 397;
Best Local Similarity 100.0%; Pred. No. 2e-130;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 217 ACTGCCTGAACATCTGAAATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGA 276
Db 12 ACTGCCTGAACATCTGAAATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGA 71
QY 277 ATTAACCTTTGAAACACCTTAAAGCTTTAAACCGCTTTAACCGGACCAACAGCATCAGAAGAT 336
Db 72 ATTAACCTTTGAAACACCTTAAAGCTTTAAACCGCTTTAACCGGACCAACAGCATCAGAAGAT 131
QY 337 AATTGCTTTACAGAAATGGGGAGGAGATCTCTGAAATCGGCCATTTCAGTCCGACTTGGATGC 396
Db 132 AATTGCTTTACAGAAATGGGGAGGAGATCTCTGAAATCGGCCATTTCAGTCCGACTTGGATGC 191
QY 397 GTTCCACTCGGGATTTCAAACATCGGCCAAAGAGTCTTGCAATACCTCTCCCGGTTTGA 456

Db 192 GTTCCACTCGGGATTTCAAACATCGGCCAAAGAGTCTTGCAATACCTCTCCCGGTTGA 251
QY 457 GAGCTGGACACCCAGGAGCGCGGTGTGTCCAG 490
Db 252 GAGCTGGACACCCAGGAGCGCGGTGTGTCCAG 285
RESULT 11
AY204568
LOCUS AY204568 3455 bp mRNA linear MAM 05-AUG-2003
DEFINITION Canis familiaris SHARPI protein (SHARPI) mRNA, complete cds.
ACCESSION AY204568
VERSION AY204568.1 GI:334339703
KEYWORDS
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 3455)
AUTHORS Kukekova, A.V., Aguirre, G.D. and Acland, G.M.
TITLE Cloning and characterization of canine SHARPI and its evaluation as a positional candidate for canine early retinal degeneration (erd)
JOURNAL Gene 312, 335-343 (2003)
REFERENCE 2 (bases 1 to 3455)
AUTHORS Kukekova, A.V., Aguirre, G.D. and Acland, G.M.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-2002) Center for Canine Genetics and Reproduction, James A. Baker Institute for Animal Health, Hungerford Hill Road, Ithaca, NY 14853-6401, USA
FEATURES
source
Location/Qualifiers
1. .3455
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/tissue type="retina"
1. .3455
/gene="SHARPI"
/gene="SHARPI"
/note="BHLH transcription factor family member; DEC2; BHLHB3"
/codon_start=1
/product="SHARPI protein"
/protein_id="AAP12465.1"
/db_xref="GI:334339704"
/translation="MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRRDSDKDT
YKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVLELTTLKHLKALTAL
TEQHQKIIIALONGERSLKSPIQSLDADFHSGFQCAKEVLQYLSFESWTPREQRCV
QLINHLHAVATQFLPTQLLTQQVPLSKGAGAPSAAPAGSAAAPCLERAGQKLEPLA
HCVPVIORTQPSAEIAAENDTDTDSYGGAEABRPDRDKGSGAGRVTIKQEPPEG
SPAPKRMKLDNRGGGGGGGGGLGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
AALLRPDAALLSILVAFGGGGAPFAQPAAPAAAAAFCPLPFYFLSPSAAAAAYVQPF
GLEKYLPAAPAAAFPLLPGIPAPAAAAAFAAAAAAFAAAAAAFAAAAAAFAAAAAA
AATLLPHEVAPPGALHPPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHP
GKEAH"

ORIGIN
Query Match 9.5%; Score 143; DB 4; Length 3455;
Best Local Similarity 100.0%; Pred. No. 2.6e-62;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 323 CAGCATCAGAAGATAATTGCTTTACAGAAATGGGAGCGGATCTCTGAAATCGCCATTG 382
Db 584 CAGCATCAGAAGATAATTGCTTTACAGAAATGGGAGCGGATCTCTGAAATCGCCATTG 643
QY 383 TCCGACTTGGATGCGTTCCTCCACTCGGGATTTCAAACATCGCCCAAGAGTCTTCAATAC 442
Db 644 TCCGACTTGGATGCGTTCCTCCACTCGGGATTTCAAACATCGCCCAAGAGTCTTCAATAC 703
QY 443 CTCTCCCGGTTTGAGAGCTGGAC 465
Db 704 CTCTCCCGGTTTGAGAGCTGGAC 726

RESULT 12
CQ687160
LOCUS CQ687160 260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 32086 from Patent WO02070737.
ACCESSION CQ687160
VERSION CQ687160.1 GI:42217526
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 32086 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source Location/Qualifiers
1. .260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 8.5%; Score 128; DB 6; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGACGAAGGAATTCCTCATTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTTAT 60
Db 105 CATGGACGAAGGAATTCCTCATTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTTAT 164
QY 61 AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGA 120
Db 165 AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGA 224
QY 121 CACCAAGG 128
Db 225 CACCAAGG 232
RESULT 13
CQ688754
LOCUS CQ688754 260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 33680 from Patent WO02070737.
ACCESSION CQ688754
VERSION CQ688754.1 GI:42221719
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 33680 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source Location/Qualifiers
1. .260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 8.5%; Score 128; DB 6; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGACGAAGGAATTCCTCATTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTTAT 60
Db 115 CATGGACGAAGGAATTCCTCATTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTTAT 174

QY 61 AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGA 120
Db 175 AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGA 234
QY 121 CACCAAGG 128
Db 235 CACCAAGG 242
RESULT 14
CQ695847
LOCUS CQ695847 260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 40773 from Patent WO02070737.
ACCESSION CQ695847
VERSION CQ695847.1 GI:42243288
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 40773 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source Location/Qualifiers
1. .260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 8.5%; Score 128; DB 6; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGACGAAGGAATTCCTCATTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTTAT 60
Db 116 CATGGACGAAGGAATTCCTCATTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTTAT 175
QY 61 AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGA 120
Db 176 AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGA 235
QY 121 CACCAAGG 128
Db 236 CACCAAGG 243
RESULT 15
CQ687963
LOCUS CQ687963 261 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 32889 from Patent WO02070737.
ACCESSION CQ687963
VERSION CQ687963.1 GI:42219444
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 32889 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source Location/Qualifiers
1. .261
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 8.5%; Score 128; DB 6; Length 261;

```
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CATGGACGAAGGAATTCCCTCATTTGCAAGAGAGACAGTACTGGAACATAGAGATTTTAT 60
      |||||||
Db      114 CATGGACGAAGGAATTCCCTCATTTGCAAGAGAGACAGTACTGGAACATAGAGATTTTAT 173
      |||||||

QY      61  AGGACTGGACTATTCCCTCTTTGTATATGTGTAAACCCAAAAGGAGCATGAAACGAGACGA 120
      |||||||
Db      174 AGGACTGGACTATTCCCTCTTTGTATATGTGTAAACCCAAAAGGAGCATGAAACGAGACGA 233
      |||||||

QY      121 CACCAAGG 128
      |||||||
Db      234 CACCAAGG 241
```

Search completed: December 17, 2004, 11:50:02
Job time : 6276.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 17, 2004, 03:13:29 ; Search time 1659.37 Seconds
(without alignments)
11518.321 Million cell updates/sec

Title: US-10-078-650-1
Perfect score: 3641
Sequence: 1 ctgcactgaaggaggagc.....ttaaaaaaaaaaaaaaaaaa 3641

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3641	100.0	3641	4	Aaf74767 Human DEC
2	3641	100.0	3641	6	Abs76400 cDNA enco
3	3641	100.0	3641	9	Acc72416 Human ova
4	3641	100.0	3641	12	Adn05013 Antipsori
5	3152	86.6	3521	10	Adc30675 Human nov
6	1583	43.5	2304	6	Abt05457 DNA of NO
7	1471	40.4	1471	10	Acd19338 cDNA enco
8	1378	37.8	1511	4	Aaf74776 Human DEC
9	629	17.3	753	4	Aal26629 Human bre
10	552	15.2	909	10	Abx74439 Human cDN
11	516	14.2	1573	12	Adq24375 Human sof
12	496	13.6	498	4	Aas26175 Human cDN
13	496	13.6	498	8	Abx73516 Human nov
14	430	11.8	430	9	Acc72415 Human cDN
15	427	11.7	628	10	Acd19339 cDNA enco
16	391	10.7	462	4	Aal16710 Human bre
17	386	10.6	447	4	Aal25553 Human bre
18	382	10.5	527	11	Adm10886 Human ova
19	382	10.5	527	12	Adj11216 Human ova
20	382	10.5	527	12	Adm43477 Human ova
21	368	10.1	896	10	Adc32475 Human nov

22	366	10.1	454	4	AAS26590	Aas26590 Human cDN
23	366	10.1	454	8	ABX73931	Abx73931 Human nov
24	355	9.8	521	5	AAH83267	Aah83267 Human ova
25	354	9.7	501	4	AAS25314	Aas25314 Human ova
26	341	9.4	580	4	AAS24664	Aas24664 Human ova
27	341	9.4	580	5	AAH83289	Aah83289 Human ova
28	339	9.3	411	4	AAL07807	Aal07807 Human bre
29	308	8.5	553	5	ADL43969	Adl43969 Human ova
30	308	8.5	684	5	ADL37585	Adl37585 Human ova
31	308	8.5	684	5	ADI72444	Adi72444 Human ova
32	296	8.1	409	4	AAS25322	Aas25322 Human ova
33	296	8.1	502	4	AAS24569	Aas24569 Human ova
34	296	8.1	502	5	AAH83185	Aah83185 Human ova
35	280	7.7	501	4	AAS24644	Aas24644 Human ova
36	274	7.5	285	4	AAS24787	Aas24787 Human ova
37	274	7.5	397	5	AAH83421	Aah83421 Human ova
38	271	7.4	322	4	AAS24447	Aas24447 Human ova
39	271	7.4	424	5	AAH83050	Aah83050 Human ova
40	266	7.3	440	6	ABL38131	Ab138131 Human col
41	246	6.8	502	3	AAC04136	Aac04136 Human sec
42	242	6.6	604	5	ADL45897	Adl45897 Human ova
43	216	5.9	216	6	ABV87290	Abv87290 Human col
44	209	5.7	312	12	ADQ20370	Adq20370 Human sof
45	194	5.3	524	5	ADL40167	Adl40167 Human ova

ALIGNMENTS

RESULT 1
AAF74767
ID AAF74767 standard; cDNA; 3641 BP.
XX
AC AAF74767;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human DEC2a encoding cDNA sequence SEQ ID NO:1.
XX
KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 135..1583
FT /*tag= a
FT /product= "DEC2a"
FT /note= "bHLH type transcription factor"

WO200114551-A1.
XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP003991.
XX
PR 19-AUG-1999; 99JP-00233286.
XX (CHUS) CHUGAI SEIYAKU KK.
XX Fujimoto K, Shin M, Kato Y;
PI WPI; 2001-202935/20.
XX P-PSDB; AAB70692.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation.
XX Claim 1; Page 48-55; 83pp; Japanese.
PS The present invention describes a basic helix loop helix (bHLH) type
XX transcription factor designated DEC2. DEC2 can be used as a tool in the

QY	2041	AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCTGTTTTAGGAGCGT	2100
Db			
Db	2041	AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCTGTTTTAGGAGCGT	2100
QY	2101	TCCACAAACAGATTGAGGCTCTTTTAGAATTGAATTTACTCTCAGTATTTTCTAATGT	2160
Db			
Db	2101	TCCACAAACAGATTGAGGCTCTTTTAGAATTGAATTTACTCTCAGTATTTTCTAATGT	2160
QY	2161	TCAGCTTTCPAAAAGGCATATATTTTCAAAGAAAGTGAGGATGCGAGTTTCTCACGTTGCA	2220
Db			
Db	2161	TCAGCTTTCTAAAAGGCATATATTTTCAAAGAAAGTGAGGATGCGAGTTTCTCACGTTGCA	2220
QY	2221	ACCTATTCTGAAGTGGTTTTAAATGGTATCTCTTAGTAACTTGCACTCGTTAAAGAAACAC	2280
Db			
Db	2221	ACCTATTCTGAAGTGGTTTTAAATGGTATCTCTTAGTAACTTGCACTCGTTAAAGAAACAC	2280
QY	2281	GGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGAAGGCCAGATCA	2340
Db			
Db	2281	GGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGAAGGCCAGATCA	2340
QY	2341	TTCCCTAGTACATTTGCTAACACCTTTTATGAGAAATTGACCATGAATTAATGGACTCATCT	2400
Db			
Db	2341	TTCCCTAGTACATTTGCTAACACCTTTTATGAGAAATTGACCATGAATTAATGGACTCATCT	2400
QY	2401	TAAATTCITCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA	2460
Db			
Db	2401	TAAATTCITCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA	2460
QY	2461	TAGATAGGTATCTATACATACACATCTCAAGTGCACTCTATTTCCCACTCTCATTAATCCAT	2520
Db			
Db	2461	TAGATAGGTATCTATACATACACATCTCAAGTGCACTCTATTTCCCACTCTCATTAATCCAT	2520
QY	2521	CATGTTCCCTAAATTTTGTAAATCTTACTGTAAAAAAAGTGCACTGAACTTCAAAAACAAA	2580
Db			
Db	2521	CATGTTCCCTAAATTTTGTAAATCTTACTGTAAAAAAAGTGCACTGAACTTCAAAAACAAA	2580
QY	2581	ACAAAAACAACAAACAACAAAAAAGTCCAACTGATATATCCTATATTTCTGTTAAAA	2640
Db			
Db	2581	ACAAAAACAACAAACAACAAAAAAGTCCAACTGATATATCCTATATTTCTGTTAAAA	2640
QY	2641	TTCAAAAGTGAACGAAAGCATTTTAAGTGGCCAGTTTTGATTGCAAAATGCTGTAAGATAT	2700
Db			
Db	2641	TTCAAAAGTGAACGAAAGCATTTTAAGTGGCCAGTTTTGATTGCAAAATGCTGTAAGATAT	2700
QY	2701	AGAATGAAGTCTGTGAGGCCCTTCCATCTCCAAGTCTATGTAATTTCTGGAGACCAAAC	2760
Db			
Db	2701	AGAATGAAGTCTGTGAGGCCCTTCCATCTCCAAGTCTATGTAATTTCTGGAGACCAAAC	2760
QY	2761	CAGATACCGAGATAATCACAAAGAAAGCTTTTAAAGGCTTAAACCAAGACCTTGCT	2820
Db			
Db	2761	CAGATACCGAGATAATCACAAAGAAAGCTTTTAAAGGCTTAAACCAAGACCTTGCT	2820
QY	2821	AGATATTTTGTGTTGTTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTA	2880
Db			
Db	2821	AGATATTTTGTGTTGTTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTA	2880
QY	2881	AGGGGTGAGACACACACAGCTGTGACTATGAGTGAGGAAAATATCTGGGTCTTTTCGTCAGT	2940
Db			
Db	2881	AGGGGTGAGACACACACAGCTGTGACTATGAGTGAGGAAAATATCTGGGTCTTTTCGTCAGT	2940
QY	2941	TTGGTGCAATTGCTGCTGCTGTTGCTACTGTTTGGCTCAAACGCTGTGTTTAAACACGCT	3000
Db			
Db	2941	TTGGTGCAATTGCTGCTGCTGTTGCTACTGTTTGGCTCAAACGCTGTGTTTAAACACGCT	3000
QY	3001	TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG	3060
Db			
Db	3001	TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG	3060
QY	3061	ATGCTGACATGCTATTTTGTAGGGAAAAAATATGTCTAATGATATTTTGAGTTAAAA	3120
Db			
Db	3061	ATGCTGACATGCTATTTTGTAGGGAAAAAATATGTCTAATGATATTTTGAGTTAAAA	3120

QY	3121	TATCTTTTGGGAGGATTTTGCTGAAAAAGTTGCACCTTTTGTTACAATGCTTATGCTTGGTA	3180
Db			
Db	3121	TATCTTTTGGGAGGATTTTGCTGAAAAAGTTGCACCTTTTGTTACAATGCTTATGCTTGGTA	3180
QY	3181	CAAGCTTATGCTGCTCTTAAATTTATTTTAAAAAAATTAATACTGCTGCTGAGAAACCAGC	3240
Db			
Db	3181	CAAGCTTATGCTGCTCTTAAATTTATTTTAAAAAAATTAATACTGCTGCTGAGAAACCAGC	3240
QY	3241	TGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAATTAACCTTTATATTTCTAGTATT	3300
Db			
Db	3241	TGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAATTAACCTTTATATTTCTAGTATT	3300
QY	3301	TTCAGCACTCCATAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT	3360
Db			
Db	3301	TTCAGCACTCCATAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT	3360
QY	3361	TCTTACTAAGGAATAAAAACTTTTAATATACGATATGATATTGTCTAATAATTAATAAAGA	3420
Db			
Db	3361	TCTTACTAAGGAATAAAAACTTTTAATATACGATATGATATTGTCTAATAATTAATAAAGA	3420
QY	3421	CATAATGGATGCTCAATTAGTTTTTAAGATATCTATAACTATAGGGATACAAATCACTACA	3480
Db			
Db	3421	CATAATGGATGCTCAATTAGTTTTTAAGATATCTATAACTATAGGGATACAAATCACTACA	3480
QY	3481	GTTCTCAGATTTACACCTTTTGTTCATGGCTTGATGTCACACATTTCCAATCTCTT	3540
Db			
Db	3481	GTTCTCAGATTTACACCTTTTGTTCATGGCTTGATGTCACACATTTCCAATCTCTT	3540
QY	3541	GCAAGCCTCCAGGCTCTGGCTTTGTCTACCTGCTCGTTCCTCAATGTATCTTAATGAAAAG	3600
Db			
Db	3541	GCAAGCCTCCAGGCTCTGGCTTTGTCTACCTGCTCGTTCCTCAATGTATCTTAATGAAAAG	3600
QY	3601	TGCAAAAGAAAAACCTACCAATTAATAAAAAAATAAAAAA 3641	
Db			
Db	3601	TGCAAAAGAAAAACCTACCAATTAATAAAAAAATAAAAAA 3641	
RESULT 2			
ABS76400			
ID	ABS76400 standard; cDNA; 3641 BP.		
XX	ABS76400;		
AC	11-DEC-2002 (first entry)		
XX	cDNA encoding human ovarian cancer marker OV4.		
DE	Human; ovarian cancer; marker; cancer; familial history; brain disorder;		
XX	central nervous system disorder; bacterial meningitis; viral meningitis;		
KW	Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;		
KW	brain herniation; inflammation; encephalitis; testicular disorder;		
KW	nontuberculous granulomatous orchitis; connective tissue disorder;		
KW	heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;		
XX	histological type; carcinogenic; ovarian cancer marker; gene; ss.		
OS	Homo sapiens.		
XX	WO200271928-A2.		
PN	19-SEP-2002.		
XX	14-MAR-2002; 2002WO-US007826.		
PF	14-MAR-2001; 2001US-0276025P.		
XX	14-MAR-2001; 2001US-0276026P.		
PR	10-AUG-2001; 2001US-0311732P.		
PR	19-SEP-2001; 2001US-0323580P.		
PR	26-SEP-2001; 2001US-0324967P.		
PR	26-SEP-2001; 2001US-0325102P.		
PR	26-SEP-2001; 2001US-0325149P.		
XX	(MILL-) MILLENNIUM PHARM INC.		
PA			
XX			

RESULT 3
ACC72416
ID ACC72416 standard; DNA; 3641 BP.
XX AC
XX ACC72416;
XX AC
DT 07-JUL-2003 (first entry)
XX Human ovarian cancer clone O168P full-length DNA.
DE Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer; gene;
XX KW ds.
KW XX
OS Homo sapiens.
XX WO2003029468-A1.
PX PD 10-APR-2003.
XX PF 02-OCT-2002; 2002WO-US031467.
XX PR 02-OCT-2001; 2001US-0327135P.
PR 30-MAY-2002; 2002US-0384531P.
XX PA (CORI-) CORIXA CORP.
XX PI Algate PA, Mannion J;
DR WPI; 2003-372001/35.
XX PT New polynucleotide and polypeptide useful for diagnosing and/or treating
PT cancer, particularly ovarian cancer, and as a vaccine.
XX PS Claim 1; Page 148-149; 169pp; English.
XX SQ The invention relates to a novel isolated polynucleotide. The
CC polynucleotides of the invention have cytostatic activity, and may have a
CC use in gene therapy, and in a vaccine. The composition and methods are
CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.
CC The composition may also be used as a vaccine to prevent cancer. The
CC present sequence is used in the exemplification of the invention
XX SQ Sequence 3641 BP; 994 A; 922 C; 836 G; 889 T; 0 U; 0 Other;
Query Match 100.0%; Score 3641; DB 9; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3641; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 CTGCACTGAAGAGGGAGAGCGGACGACAGAGACTGGAGACGCACAGATCCCCCAAGGTCTC 60
DB 1 CTGCACTGAAGAGGGAGAGCGGACGACAGAGACTGGAGACGCACAGATCCCCCAAGGTCTC 60
QY 61 CCAAGCCTACCGTCCCACAGATTATTGTACAGAGCCCCAAAATCGAAACAGAGGAAACG 120
DB 61 CCAAGCCTACCGTCCCACAGATTATTGTACAGAGCCCCAAAATCGAAACAGAGGAAACG 120
QY 121 AACAGCAGTTGAACATGGACGAGGAATTCCTCATTTTGTAAGAGACAGTACTGGAAC 180
DB 121 AACAGCAGTTGAACATGGACGAGGAATTCCTCATTTTGTAAGAGACAGTACTGGAAC 180
QY 181 ATAGAGATTTTAGGACTGGACTATTCCTCTTTTGTATATGTGTAAACCCCAAAGGACA 240
DB 181 ATAGAGATTTTAGGACTGGACTATTCCTCTTTTGTATATGTGTAAACCCCAAAGGACA 240
QY 241 TGAAACGAGACGACACCCAAGGATACTCAAAATTACCGCACAGATTATAGAAAAAGAAA 300
DB 241 TGAAACGAGACGACACCCAAGGATACTCAAAATTACCGCACAGATTATAGAAAAAGAAA 300
QY 301 GAAGAGACCGGAATTAATGAATGCATTGCTCAGCTGAAAGATTTTACTGCCTGAACATCTGA 360
DB 301 GAAGAGACCGGAATTAATGAATGCATTGCTCAGCTGAAAGATTTTACTGCCTGAACATCTGA 360
QY 361 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGATTAACTTTGAAACACT 420

RESULT 4
ID ADN05013 standard; cDNA; 3641 BP.
XX
AC ADN05013;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #721.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-305105/28.
DR P-PSDB; ADN05014.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 1; SEQ ID NO 1407; 3069pp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 3641 BP; 994 A; 922 C; 836 G; 889 T; 0 U; 0 Other;
Query Match 100.0%; Score 3641; DB 12; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCACTGAAGAGGGAGAGCGGAGAGAGAGACTGGAGAGCGCACAGATCCCCCAAGGTCTC 60
Db 1 CTGCACTGAAGAGGGAGAGCGGAGAGAGAGACTGGAGAGCGCACAGATCCCCCAAGGTCTC 60
QY 61 CCAAGCTACCGTCCACAGATTATTGTACAGAGCCCCCAAAATCGAAACAGAGGAAACG 120
Db 61 CCAAGCTACCGTCCACAGATTATTGTACAGAGCCCCCAAAATCGAAACAGAGGAAACG 120
QY 121 AACAGCAGTTGAACATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTACTGGAAC 180
Db 121 AACAGCAGTTGAACATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTACTGGAAC 180
QY 181 ATAGAGATTTTATAGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCA 240
Db 181 ATAGAGATTTTATAGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCA 240
QY 241 TGAACGAGACGACACCAAGGATACCTACAAATTACCGCACAGATTATAGAAAGAAAA 300
Db 241 TGAACGAGACGACACCAAGGATACCTACAAATTACCGCACAGATTATAGAAAGAAAA 300
QY 301 GAAGAGACCGAATTAAATGAATGCATTGCTCAGCTGAAGATTACTGCCTGAACATCTGA 360
Db 301 GAAGAGACCGAATTAAATGAATGCATTGCTCAGCTGAAGATTACTGCCTGAACATCTGA 360

QY 361 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTTGGAATTAACCTTGAACACT 420
Db 361 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTTGGAATTAACCTTGAACACT 420
QY 421 TAAAAGCTTTAACCGCCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATG 480
Db 421 TAAAAGCTTTAACCGCCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATG 480
QY 481 GGGAGCGATCTCTGAATCGCCCATTCAGTCCGACTTGGATCGTTCCTCCTCGGATTTC 540
Db 481 GGGAGCGATCTCTGAATCGCCCATTCAGTCCGACTTGGATCGTTCCTCCTCGGATTTC 540
QY 541 AAACATGCGCCAAAGAAAGTCTTGCAATACCTCTCCCGTTTGGAGCTGGACACCCAGGG 600
Db 541 AAACATGCGCCAAAGAAAGTCTTGCAATACCTCTCCCGTTTGGAGCTGGACACCCAGGG 600
QY 601 AGCCCGGTGTGTCCAGCTGATCAACCACTTGACGCCGTGGCCACCCAGTCTTGTGCCCA 660
Db 601 AGCCCGGTGTGTCCAGCTGATCAACCACTTGACGCCGTGGCCACCCAGTCTTGTGCCCA 660
QY 661 CCCGCGAGTGTGACTCAACAGGTCCCTCTGAGAAAAGGACCCGGCGTCCCTCGGCCG 720
Db 661 CCCGCGAGTGTGACTCAACAGGTCCCTCTGAGAAAAGGACCCGGCGTCCCTCGGCCG 720
QY 721 CCGGTCCGCGCCCGCTGCTGGAGCGCGCGGCGAGAAGCTGGAGCCCTCGCCT 780
Db 721 CCGGTCCGCGCCCGCTGCTGGAGCGCGCGGCGAGAAGCTGGAGCCCTCGCCT 780
QY 781 ACTGGTCCCGCTCATCCAGCGGACTCAGCCAGCGCCGAGCTCGCCGCGAGAACGACA 840
Db 781 ACTGGTCCCGCTCATCCAGCGGACTCAGCCAGCGCCGAGCTCGCCGCGAGAACGACA 840
QY 841 CGGACACCGACAGCGGTACGGCGGGAAGCCGAGSCCGCGGACCGCGAGAAAGGCA 900
Db 841 CGGACACCGACAGCGGTACGGCGGGAAGCCGAGSCCGCGGACCGCGAGAAAGGCA 900
QY 901 AAGSCGCGGGGCGAGCCGCTCACCATCAAGCAGAGCCCTCCCGGGAGGACTCGCCGG 960
Db 901 AAGSCGCGGGGCGAGCCGCTCACCATCAAGCAGAGCCCTCCCGGGAGGACTCGCCGG 960
QY 961 CGCCCAAGAGGATGAAGCTGGATTCCCGCGCGCGGCGAGCGCGCGCGCGGGGGCG 1020
Db 961 CGCCCAAGAGGATGAAGCTGGATTCCCGCGCGCGGCGAGCGCGCGCGCGGGGGCG 1020
QY 1021 GCGCGGCGGCGGCGAGCCGCTTCTGGGGCCCGACCCCTGCGCCGCGCGCGCGCTGC 1080
Db 1021 GCGCGGCGGCGGCGAGCCGCTTCTGGGGCCCGACCCCTGCGCCGCGCGCGCGCTGC 1080
QY 1081 TGAGACCCGACGCGCCCTGCTCAGCTCGTGTGGTGGCTTGGCGGAGGCGGAGGCGCGC 1140
Db 1081 TGAGACCCGACGCGCCCTGCTCAGCTCGTGTGGTGGCTTGGCGGAGGCGGAGGCGCGC 1140
QY 1141 CCTTCCCGCAGCCCG 1200
Db 1141 CCTTCCCGCAGCCCG 1200
QY 1201 CTTTGTGAGCTGCGCGCTACGTGACGCCCTTCTGAGCAAGAGCGGCTTGAGAGAGTATC 1260
Db 1201 CTTTGTGAGCTGCGCGCTACGTGACGCCCTTCTGAGCAAGAGCGGCTTGAGAGAGTATC 1260
QY 1261 TGTACCCGCGCGGCTGCGCGCCCGCTTCCGCTCCGCTGCTATACCCCGCGCATCCCCCG 1320
Db 1261 TGTACCCGCGCGGCTGCGCGCCCGCTTCCGCTCCGCTGCTATACCCCGCGCATCCCCCG 1320
QY 1321 CGGACGCGCGGCGAGCCG 1380
Db 1321 CGGACGCGCGGCGAGCCG 1380
QY 1381 CCTCGGTGTGTGCGCCCGCTTCCCGAGAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CCTCGGTGTGTGCGCCCGCTTCCCGAGAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 ACGAGGTGGCGCCCTTGGGGGCGCGCACCCCCAGCACCCCGCACGGCCCGCACCCACCTGC 1500

Db 1441 ACGAGGTGGCGCCCTTGGGGCGCCGACCCCGACCCGACGGCGCACCCACCTGC 1500
QY 1501 CCTTCGCGGGCCCCCGAGCGCGGGAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 1560
Db 1501 CCTTCGCGGGCCCCCGAGCGCGGGAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 1560
QY 1561 AGCCAGGAAAGGAGCTCCCTGAAATCCTTGGCTCCCGAAGGACGGAGGTTCAAGCAGAGT 1620
Db 1561 AGCCAGGAAAGGAGCTCCCTGAAATCCTTGGCTCCCGAAGGACGGAGGTTCAAGCAGAGT 1620
QY 1621 GAGAAAGTTAAATACCCCTTAAGGAGGTTCAAGCAGAGTGAGAAAGTTAAATACCCCTTAAG 1680
Db 1621 GAGAAAGTTAAATACCCCTTAAGGAGGTTCAAGCAGAGTGAGAAAGTTAAATACCCCTTAAG 1680
QY 1681 GTCTTTAAGGAGGAAAGTGTAATAGATGCACGACAGGCATAAACAGAAACAACAAACAG 1740
Db 1681 GTCTTTAAGGAGGAAAGTGTAATAGATGCACGACAGGCATAAACAGAAACAACAAACAG 1740
QY 1741 GTGTTAAGTGTACATTCGGAGTTCCTGTTTGGCTCATCCCGCACCCACCCCTCCACA 1800
Db 1741 GTGTTAAGTGTACATTCGGAGTTCCTGTTTGGCTCATCCCGCACCCACCCCTCCACA 1800
QY 1801 CACTAACATCCCTTCTTCCCCCACCAGCTGTAAAGATCCTATGCGAAAGACACTGGC 1860
Db 1801 CACTAACATCCCTTCTTCCCCCACCAGCTGTAAAGATCCTATGCGAAAGACACTGGC 1860
QY 1861 TCTTTTTTTAATCCCCCAATAAATTTTGCCTCTTTTAGGCCATGTTCCATTATCTCT 1920
Db 1861 TCTTTTTTTAATCCCCCAATAAATTTTGCCTCTTTTAGGCCATGTTCCATTATCTCT 1920
QY 1921 TAAATTTGGAACCTTAATTCGAGAGGAAGTAAGAGGGTCTGTTCTGTGGCTGAGCTAGGT 1980
Db 1921 TAAATTTGGAACCTTAATTCGAGAGGAAGTAAGAGGGTCTGTTCTGTGGCTGAGCTAGGT 1980
QY 1981 GAACCCCGGGTAGGGAAAGATGTTAACACCTTTGACGCTCTTTGGAGTTGACATGGAAC 2040
Db 1981 GAACCCCGGGTAGGGAAAGATGTTAACACCTTTGACGCTCTTTGGAGTTGACATGGAAC 2040
QY 2041 AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCCTGTTTtaggagggct 2100
Db 2041 AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCCTGTTTtaggagggct 2100
QY 2101 TCCACAAACAGATTGAGGCTCTTTTtagaattgaaatttactcttcagtatttttctaattgt 2160
Db 2101 TCCACAAACAGATTGAGGCTCTTTTtagaattgaaatttactcttcagtatttttctaattgt 2160
QY 2161 TCAGCTTTCAAAGGCATATATTTTCAAAGAAAGTGAGATGCAGTTTCTCACGTTGCA 2220
Db 2161 TCAGCTTTCAAAGGCATATATTTTCAAAGAAAGTGAGATGCAGTTTCTCACGTTGCA 2220
QY 2221 ACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACCTTGCACTCGTTAAAGAAACAC 2280
Db 2221 ACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACCTTGCACTCGTTAAAGAAACAC 2280
QY 2281 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGATCA 2340
Db 2281 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGATCA 2340
QY 2341 TTCCCTAGTACATTTGCTAACACCTTTTATGAGAAATTGACCATGAATTAATGGACTCATCT 2400
Db 2341 TTCCCTAGTACATTTGCTAACACCTTTTATGAGAAATTGACCATGAATTAATGGACTCATCT 2400
QY 2401 TAATTTCTCTAAGTCCATATATATAGATAGATATCTATCTGTACAGATTTTCTATTATCCA 2460
Db 2401 TAATTTCTCTAAGTCCATATATATAGATAGATATCTATCTGTACAGATTTTCTATTATCCA 2460
QY 2461 TAGATAGTATCTATACATACACATCTCAAGTGCATCTATTCCTACTCTCATTAATCCAT 2520
Db 2461 TAGATAGTATCTATACATACACATCTCAAGTGCATCTATTCCTACTCTCATTAATCCAT 2520
QY 2521 CATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAAGTGCACCTGAACTTCAAACAAA 2580

Db 2521 CATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAAGTGCACCTTCAAACAAA 2580
QY 2581 ACAAAAAACAACAACAACAAAAACAAGTCCAACTGATATATCTCTATATTTCTGTTAAAA 2640
Db 2581 ACAAAAAACAACAACAACAAAAACAAGTCCAACTGATATATCTCTATATTTCTGTTAAAA 2640
QY 2641 TTCAAAAGTGAACGAAACCATTTAACTGGCCAGTTTGTATTGCAAAATGCTGTAAAGATAT 2700
Db 2641 TTCAAAAGTGAACGAAACCATTTAACTGGCCAGTTTGTATTGCAAAATGCTGTAAAGATAT 2700
QY 2701 AGAATGAAGTCTCTGTGAGGCCCTTCTATCTCCAAGTCTATGTATTTTCTGGAGACCAAC 2760
Db 2701 AGAATGAAGTCTCTGTGAGGCCCTTCTATCTCCAAGTCTATGTATTTTCTGGAGACCAAC 2760
QY 2761 CAGATACCAGATAATCACAAAGAAAGCTTTTTTAATAAGCTTAAACCAAGACCTTGCT 2820
Db 2761 CAGATACCAGATAATCACAAAGAAAGCTTTTTTAATAAGGCTTAAACCAAGACCTTGCT 2820
QY 2821 AGATATTTTGTAGTTTGTGCCAAGGTAGCTGTGAGAAATCTCACTTGGATGTTATGTA 2880
Db 2821 AGATATTTTGTAGTTTGTGCCAAGGTAGCTGTGAGAAATCTCACTTGGATGTTATGTA 2880
QY 2881 AGGGGTGAGACACAAACAGTCTGACTATGAGTGAGGAAAAATATCTGGTCTTTTCGTGAGT 2940
Db 2881 AGGGGTGAGACACAAACAGTCTGACTATGAGTGAGGAAAAATATCTGGTCTTTTCGTGAGT 2940
QY 2941 TTGGTGCATTTGCTGCTGCTGTTGCTACTGTTGCCTCAAACGCTGTGTTTAAACCAACGT 3000
Db 2941 TTGGTGCATTTGCTGCTGCTGTTGCTACTGTTGCCTCAAACGCTGTGTTTAAACCAACGT 3000
QY 3001 TAAACTCTTAGCCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG 3060
Db 3001 TAAACTCTTAGCCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG 3060
QY 3061 ATGTCTGACATGCTATTTTGTAGGGAGAAAAATATGTCTAATGATATTTTGAGTTAAAA 3120
Db 3061 ATGTCTGACATGCTATTTTGTAGGGAGAAAAATATGTCTAATGATATTTTGAGTTAAAA 3120
QY 3121 TATCTTTTGGGGAGGATTTGCTGAAAAAGTTGCACCTTTTGTACAATGCTTATGCTTGGTA 3180
Db 3121 TATCTTTTGGGGAGGATTTGCTGAAAAAGTTGCACCTTTTGTACAATGCTTATGCTTGGTA 3180
QY 3181 CAAGCTTATGCTGCTCTTAAATTTTAAAAAAATTAATACTGTCTGTGAGAAACCAGC 3240
Db 3181 CAAGCTTATGCTGCTCTTAAATTTTAAAAAAATTAATACTGTCTGTGAGAAACCAGC 3240
QY 3241 TGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAATTAACCTTTATATCTAGTATT 3300
Db 3241 TGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAATTAACCTTTATATCTAGTATT 3300
QY 3301 TTCAGCACTCCATAAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT 3360
Db 3301 TTCAGCACTCCATAAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT 3360
QY 3361 TCTTACTAAGGAATAAAAACTTTTAAATATACGATATGATATTGTCTAATAATTAATAAAGA 3420
Db 3361 TCTTACTAAGGAATAAAAACTTTTAAATATACGATATGATATTGTCTAATAATTAATAAAGA 3420
QY 3421 CATAATGGATGCTCAATTAGTTTAAAGATATCTATAACTATAGGATACAAATCACTACA 3480
Db 3421 CATAATGGATGCTCAATTAGTTTAAAGATATCTATAACTATAGGATACAAATCACTACA 3480
QY 3481 GTTCTCAGATTTACACCTTTTGTGTCATTTGGCTTGCCTTCCCAATGTATCTTAATGAAAAG 3540
Db 3481 GTTCTCAGATTTACACCTTTTGTGTCATTTGGCTTGCCTTGCATTCACACTTTCCAATCTCT 3540
QY 3541 GCAAGCCTCCAGGCTCTGGCTTTTGTCTACCTGCTGCTTCCCAATGTATCTTAATGAAAAG 3600
Db 3541 GCAAGCCTCCAGGCTCTGGCTTTTGTCTACCTGCTGCTTCCCAATGTATCTTAATGAAAAG 3600
QY 3601 TGCAAAAGHAAAAACCTACCAATTAATAAAAAA 3641
Db 3601 TGCAAAAGHAAAAACCTACCAATTAATAAAAAA 3641

QY 901 AAGCGCGGGGCGAGCCGCGTCACCATCAACAGGAGCCCTCCCGGGAGGACTCGCCGG 960
Db 1017 AAGCGCGGGGCGAGCCGCGTCACCATCAACAGGAGCCCTCCCGGGAGGACTCGCCGG 1076
QY 961 CGCCCAAGAGGATGAAGCTGGATTCCCGCGGCGGCGGAGCGGCGGCGGCCCGGGGGCG 1020
Db 1077 CGCCCAAGAGGATGAAGCTGGATTCCCGCGGCGGCGGAGCGGCGGCGGCCCGGGGGCG 1136
QY 1021 GCGCGGCGGCGGCGGAGCCGCGCTTCTGGGGCCCGACCCCTGCCGCGCGCGCGCGCTGC 1080
Db 1137 GCGCGGCGGCGGCGGAGCCGCGCTTCTGGGGCCCGACCCCTGCCGCGCGCGCGCGCTGC 1196
QY 1081 TGAGACCCGACGCGCGCCCTGCTCAGCTCGTGGTGGCGTTCCGGCGAGCGGAGCGCGC 1140
Db 1197 TGAGACCCGACGCGCGCCCTGCTCAGCTCGTGGTGGCGTTCCGGCGAGCGGAGCGCGC 1256
QY 1141 CCTTCCCGCAGCCGCGCGCCGCGCTTCTGCGCGCCCTTCTGCGCTGCGCTTCTCTCTCGC 1200
Db 1257 CCTTCCCGCAGCCGCGCGCCGCGCTTCTGCGCGCCCTTCTGCGCTGCGCTTCTCTCTCGC 1316
QY 1201 CTTCTGAGCTGCGCGCCCTACGTGACGCCCTTCTTGACAAAGAGCGCCCTGGAGAAATC 1260
Db 1317 CTTCTGAGCTGCGCGCCCTACGTGACGCCCTTCTTGACAAAGAGCGCCCTGGAGAAATC 1376
QY 1261 TGTACCCGCGGCGGCTGCCGCGCCGCTTCCCGCTGCTATACCCCGCATCCCGCCCGG 1320
Db 1377 TGTACCCGCGGCGGCTGCCGCGCCGCTTCCCGCTGCTATACCCCGCATCCCGCCCGG 1436
QY 1321 CGGCAGCCGCGCAGCCGCGCGCGCGCTGCCGCGCGCGCGCGCGCGGCTTCCCTGCTGT 1380
Db 1437 CGGCAGCCGCGCAGCCGCGCGCGCGCTGCCGCGCGCGCGCGCGCGGCTTCCCTGCTGT 1496
QY 1381 CCTCGGTGTTGTCGCCCCCTCCCGAGAGCGGGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1497 CCTCGGTGTTGTCGCCCCCTCCCGAGAGCGGGCGCGCGCGCGCGCGCGCGCGCG 1556
QY 1441 ACGAGGTGGCGCCCTTGGGGCGCGCACCAGCAACCGGACGGCGCGCGCGCGCGCG 1500
Db 1557 ACGAGGTGGCGCCCTTGGGGCGCGCACCAGCAACCGGACGGCGCGCGCGCGCGCG 1616
QY 1501 CCTTCGCGCGCGCGCGAGCCGCGGGAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 1560
Db 1617 CCTTCGCGCGCGCGCGAGCCGCGGGAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 1676
QY 1561 AGCCAGGAAGGAAGCTCCCTGAATCCTTGCCTCCGAAAGGACGGAGGTTCAAGCAGAT 1620
Db 1677 AGCCAGGAAGGAAGCTCCCTGAATCCTTGCCTCCGAAAGGACGGAGGTTCAAGCAGAT 1736
QY 1621 GAGAAGTTAAATAACCTTAAGGAGGTTCAAGCAGAGTGAGAAGTTAAATAACCTTAAG 1680
Db 1737 GAGAAGTTAAATAACCTTAAGGAGGTTCAAGCAGAGTGAGAAGTTAAATAACCTTAAG 1796
QY 1681 GTCTTTAAGGAGGAAGTGTAATAGATGCACGACAGGCATAAACAAGAAACAACAACAG 1740
Db 1797 GTCTTTAAGGAGGAAGTGTAATAGATGCACGACAGGCATAAACAAGAAACAACAACAG 1856
QY 1741 GTGTTATGTGPACATTCGGAGTTCCTGTTTGTCTCATCCCGACACCCACCCCTCCACA 1800
Db 1857 GTGTTATGTGPACATTCGGAGTTCCTGTTTGTCTCATCCCGACACCCACCCCTCCACA 1916
QY 1801 CACTAACATCCCTTCTTCCCGCCACAGCTGTAAAGATCCTATGCGAAAGACACTGGC 1860
Db 1917 CACTAACATCCCTTCTTCCCGCCACAGCTGTAAAGATCCTATGCGAAAGACACTGGC 1976
QY 1861 TCTTTTTTTTAAATCCCGCCAAATAAATTTGCCCGCTTTAGGCCATGTTCCATATCTCT 1920
Db 1977 TCTTTTTTTTAAATCCCGCCAAATAAATTTGCCCGCTTTAGGCCATGTTCCATATCTCT 2036
QY 1921 TAAATTTGAAACCTAATTCGAGAGGAAGTAAGAAGGCTGTGTTCTGTGGCTGAGCTAGGT 1980
Db 2037 TAAATTTGAAACCTAATTCGAGAGGAAGTAAGAAGGCTGTGTTCTGTGGCTGAGCTAGGT 2096
QY 1981 GAACCCCGGGGTAGGGAAAGATGTTAACACCTTTGACGCTCTTTGGAGTTGACATGGAAC 2040

Db 2097 GAACCCCGGGGTAGGGAAAGATGTTAACACCTTTGACGCTCTTTGGAGTTGACATGGAAC 2156
QY 2041 AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCCCTGTTTAGGAGGCGT 2100
Db 2157 AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCCCTGTTTAGGAGGCGT 2216
QY 2101 TCCACAAACAGATTGAGGCTCTTTTGTAGAAATTTACTCTTTCAGTATTTTCTAATGT 2160
Db 2217 TCCACAAACAGATTGAGGCTC-TTTTAGAATTGAATTTACTCTTTCAGTATTTTCTAATGT 2275
QY 2161 TCAGCTTTCTAAAGGCATATATTTTTCAAAGAGTGAGGATGCAGTTTCTCAGCTTGCA 2220
Db 2276 TCAGCTTTCTAAAGGCATATATTTTTCAAAGAGTGAGGATGCAGTTTCTCAGCTTGCA 2335
QY 2221 ACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAACAC 2280
Db 2336 ACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAACAC 2395
QY 2281 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGAAAGGAGATGGATGAGAGGCCAGAAATCA 2340
Db 2396 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGAAAGGAGATGGATGAGAGGCCAGAAATCA 2455
QY 2341 TTCTTAGTACATTGCTAACACCTTTATTGAGAAATTGACCATGAATTAATGGACTCATCT 2400
Db 2456 TTCTTAGTACATTGCTAACACCTTTATTGAGAAATTGACCATGAATTAATGGACTCATCT 2515
QY 2401 TAAATTTCTTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTTATTTATCCA 2460
Db 2516 TAAATTTCTTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTTATTTATCCA 2575
QY 2461 TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCACCTCATTAATCCAT 2520
Db 2576 TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCACCTCATTAATCCAT 2635
QY 2521 CATGTTCCCTAAATTTTGTAACTTACTGTAAAAAAAAGTGCACCTGAACTTCAAAACAAA 2580
Db 2636 CATGTTCCCTAAATTTTGTAACTTACTGTAAAAAAAAGTGCACCTGAACTTCAAAACAAA 2695
QY 2581 ACAAAAAACAACAACAACAACAACAAGTCCAAACTGATATATCTTATATCTGTTAAAA 2640
Db 2696 ACAAAAAACAACAACAACAACAAGTCCAAACTGATATATCTTATATCTGTTAAAA 2755
QY 2641 TTCAAAGGTGAACGAAAGCATTTAACTGGCCAGTTTGTGTTGCAAAATGCTGTAAGATAT 2700
Db 2756 TTCAAAGGTGAACGAAAGCATTTAACTGGCCAGTTTGTGTTGCAAAATGCTGTAAGATAT 2815
QY 2701 AGAATGAAGTCTGTGAGGCGCTTCTATCTCCAAGTCTATGTATTTCTGGAGACCAAC 2760
Db 2816 AGAATGAAGTCTGTGAGGCGCTTCTATCTCCAAGTCTATGTATTTCTGGAGACCAAC 2875
QY 2761 CAGATACCAGATAATCACAAAGAAAGCTTTTAAAGGCTTAAAGCTTAAAGCTTGTCT 2820
Db 2876 CAGATACCAGATAATCACAAAGAAAGCTTTTAAAGGCTTAAAGCTTAAAGCTTGTCT 2935
QY 2821 AGATATTTTGTAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTGGATGTTATGTA 2880
Db 2936 AGATATTTTGTAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTGGATGTTATGTA 2995
QY 2881 AGGGGTGAGACACAACAGTCTGACTATGAGTGAGGAAATATCTGGGTCTTTTCGTGAGT 2940
Db 2996 AGGGGTGAGACACAACAGTCTGACTATGAGTGAGGAAATATCTGGGTCTTTTCGTGAGT 3055
QY 2941 TTGGTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
Db 3056 TTGGTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3115
QY 3001 TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG 3060
Db 3116 TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG 3175
QY 3061 ATGCTGACATGCTATTTTGTAGGAGAAATATGTGCTAATGATATTTTGAGTTAAAA 3120

Db 3176 ATGCTGACATGCTATTTTGTAGGGAGAAAATATGCTGCTAATGATATTTTGAGTTAAAA 3235
Qy 3121 TATCTTTTGGGGAGGATTGCTGAAAAAGTTGCACCTTTTGTACAAATGCTTATGCTTGGTA 3180
Db 3236 TATCTTTTGGGGAGGATTGCTGAAAAAGTTGCACCTTTTGTACAAATGCTTATGCTTGGTA 3295
Qy 3181 CAAGCTTATGCTGCTCTTAATATTATTTTAAAAAAATTAATACTGCTGCTGAGAAAACGAGC 3240
Db 3296 CAAGCTTATGCTGCTCTTAATATTATTTTAAAAAAATTAATACTGCTGCTGAGAAAACGAGC 3354
Qy 3241 TGGTTAGAAAAGTTTAGTATGTGACGATAAACTAGAAATACCTTTATATCTAGTATT 3300
Db 3355 TGGTTAGAAAAGTTTAGTATGTGACGATAAACTAGAAATACCTTTATATCTAGTATT 3414
Qy 3301 TTCAGCACTCCATAAAATCTTATTAACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT 3360
Db 3415 TTCAGCACTCCATAAAATCTTATTAACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT 3474
Qy 3361 TCTTACTAAGGAATAAAAACTTTAATATACGATA 3394
Db 3475 TCTTACTAAGGAATAAAAACTTTAATATACGATA 3508

RESULT 6

ABT05457

ID ABT05457 standard; DNA; 2304 BP.

XX

AC ABT05457;

XX

DT 11-OCT-2002 (first entry)

XX

DE DNA of NOVX 5 SEQ ID No 9.

XX

KW Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipaemic;
KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
KW tranquiliser; neuroleptic; antidiabetic; antiulcer; antiinflammatory;
KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
KW vesicular transport; cystic fibrosis; gastrointestinal disorder;
KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
KW gene therapy; gene; ds.

XX

OS Unidentified.

XX

PN WO200246409-A2.

XX

PD 13-JUN-2002.

XX

PF 06-DEC-2001; 2001WO-US046586.

XX

PR 06-DEC-2000; 2000US-0251660P.

PR

PR 12-DEC-2000; 2000US-0255029P.

PR

PR 08-JAN-2001; 2001US-0260326P.

PR

PR 24-JAN-2001; 2001US-0263800P.

PR

PR 20-FEB-2001; 2001US-0269942P.

PR

PR 24-APR-2001; 2001US-0286183P.

PR

PR 20-AUG-2001; 2001US-0313627P.

PR

PR 12-SEP-2001; 2001US-0318712P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Guo X, Li L, Patturajan M, Shimkets RA, Casman SJ, Malyankar UM;

PI Tchernev VT, Vernet CAM, Spytek KA, Shenoy SG, Alsobrook JP;

PI Edinger S, Peyman JA, Stone DJ, Ellefman K, Gangolli EA, Boldog FL;

PI Colman SD, Eisen AJ, Liu X, Padigaru M, Spaderna SK, Zerhusen BD;

XX WPI; 2002-547774/58.

DR P-PSDB; ABJ04644.

DR

XX Novel isolated polypeptide, designated NOVX, useful for treating or
PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT metabolic, neurodegenerative, immune and hematopoietic disorders.
XX
PS Claim 9; Page 43-44; 421pp; English.
XX
CC The invention relates to an isolated polypeptide, designated NOVX,
CC comprising a sequence fully defined in the specification. The isolated
CC protein, its encoding polynucleotide or an antibody created from the
CC protein is useful in the manufacture of a medicament for treating a
CC syndrome associated with a human disease, preferably a NOVX-associated
CC disorder, or for treating or preventing a NOVX-associated disorder in a
CC subject, preferably human. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are also useful
CC for treating or preventing metabolic disorders, diabetes, obesity,
CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
CC disease, Parkinson's disorder, immune disorders, haematopoietic
CC disorders, and various dyslipidaemias, metabolic disturbances associated
CC with obesity, the metabolic syndrome X, wasting disorders associated with
CC chronic diseases, and cancer. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are useful for
CC treating or preventing neurological disorders such as epilepsy, stroke,
CC mental disorders including mood, anxiety, schizophrenic disorders,
CC disorders of vesicular transport such as cystic fibrosis, diabetes
CC mellitus, goiter, gastrointestinal disorders including ulcerative
CC colitis, other conditions associated with abnormal vesicle trafficking
CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
CC arthritis. A cell comprising the vector of the invention is useful for
CC producing non-human transgenic animals. The polynucleotide of the
CC invention can be used to treat disorders by gene therapy. This
CC polynucleotide sequence represents the DNA encoding one of the isolated
CC NOVX proteins of the invention

XX SQ Sequence 2304 BP; 468 A; 784 C; 689 G; 363 T; 0 U; 0 Other;

Query Match 43.5%; Score 1583; DB 6; Length 2304;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACTGAAGAGGGAGAGCGAGAGAGAGACTGGAGACGACAGATCCCCCAAGGTCTC 60
Db 722 CTGCACTGAAGAGGGAGAGCGAGAGAGAGACTGGAGACGACAGATCCCCCAAGGTCTC 781
Qy 61 CCAAGCCTACCGTCCACAGATTATTGTACAGAGCCCCAAAAATCGAAACAGAGGAAACG 120
Db 782 CCAAGCCTACCGTCCACAGATTATTGTACAGAGCCCCAAAAATCGAAACAGAGGAAACG 841
Qy 121 AACAGCAGTTGAACATGGACGAAGGAATTCTCTATTGCAAGAGAGACAGTTACTGGAAC 180
Db 842 AACAGCAGTTGAACATGGACGAAGGAATTCTCTATTGCAAGAGAGACAGTTACTGGAAC 901
Qy 181 ATAGAGATTTTATAGGACTGGACTATTCTCTTTTGTATATGTGTAAACCCCAAGAGGACA 240
Db 902 ATAGAGATTTTATAGGACTGGACTATTCTCTTTTGTATATGTGTAAACCCCAAGAGGACA 961
Qy 241 TGAACGAGACGACACCAAGGATACCTACAAATTACCAGACAGATTAAATAGAAAAGAAA 300
Db 962 TGAACGAGACGACACCAAGGATACCTACAAATTACCAGACAGATTAAATAGAAAAGAAA 1021
Qy 301 GAAGAGACCGAATTAAATGAATGCATTGCTCAGCTGAAAGATTTACTGCCTGACATCTGA 360
Db 1022 GAAGAGACCGAATTAAATGAATGCATTGCTCAGCTGAAAGATTTACTGCCTGACATCTGA 1081
Qy 361 AATTGACAACTCTGGGACATCTGGAGAAAAGCTGTAGTCTTGGAAATTAACTTTGAAACACT 420
Db 1082 AATTGACAACTCTGGGACATCTGGAGAAAAGCTGTAGTCTTGGAAATTAACTTTGAAACACT 1141
Qy 421 TAAAAGCTTTAAACCGCTTAAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAAATG 480
Db 1142 TAAAAGCTTTAAACCGCTTAAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAAATG 1201
Qy 481 GGGAGCGATCTCTGAAATCGCCCATTCAGTCCGACTTGGATCGGTTCCACTCGGGAATTC 540

Db 1202 GGGAGCGATCTCTGAAATCGCCCATTCAGTCCGACTTGGATCGTTCCACTCGGATTC 1261

QY 541 AAACATGCGCCAAAGAAGTCTTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 600

Db 1262 AAACATGCGCCAAAGAAGTCTTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 1321

QY 601 AGCCGCGGTGTGTCAGCTGATCAACCACTTGACCCCGCTGGCCACCCAGTTCTTGCCCA 660

Db 1322 AGCCGCGGTGTGTCAGCTGATCAACCACTTGACCCCGCTGGCCACCCAGTTCTTGCCCA 1381

QY 661 CCCCAGCTGTTGACTCAACAGGTCCCTCTGAGCAAGGCAACCGCGCTCCCTCGGCCG 720

Db 1382 CCCCAGCTGTTGACTCAACAGGTCCCTCTGAGCAAGGCAACCGCGCTCCCTCGGCCG 1441

QY 721 CCGGTCCGCGCGCCCTCTGCTGGAGCGCGGGGCAAGCTGGAGCCCTCGCCT 780

Db 1442 CCGGTCCGCGCGCCCTCTGCTGGAGCGCGGGGCAAGCTGGAGCCCTCGCCT 1501

QY 781 ACTGCTGCCGCTCATCCAGCGGACTCAGCCAGCGCGGAGCTCGCGCGGAGAACGACA 840

Db 1502 ACTGCTGCCGCTCATCCAGCGGACTCAGCCAGCGCGGAGCTCGCGCGGAGAACGACA 1561

QY 841 CGGACACCGACGCGGCTACGGCGGCGAAGCGAGGCGCGCGCGCGGAGAACGACA 900

Db 1562 CGGACACCGACGCGGCTACGGCGGCGAAGCGAGGCGCGCGCGGAGAACGACA 1621

QY 901 AAGCGCGGGGCGAGCCCGCTACCATCAAGCAGGAGCCTCCCGGGGAGGACTCGCCGG 960

Db 1622 AAGCGCGGGGCGAGCCCGCTACCATCAAGCAGGAGCCTCCCGGGGAGGACTCGCCGG 1681

QY 961 CGCCCAAGAGGATGAAGCTGGATTCCCGGGCGGCGGAGCGCGCGCGCGCGGGGGG 1020

Db 1682 CGCCCAAGAGGATGAAGCTGGATTCCCGGGCGGCGGAGCGCGCGCGCGCGGGGGG 1741

QY 1021 GCGCGGCGGCGGCGGCGGCTTCTGGGGCCCGACCCCTGCCCGCGCGCGCGCTGC 1080

Db 1742 GCGCGGCGGCGGCGGCGGCTTCTGGGGCCCGACCCCTGCCCGCGCGCGCGCTGC 1801

QY 1081 TGAGACCCGACGCGCCCTGTCTCAGCTCGTGGTGGCGTTCCGGCGGAGCGGAGCGCGC 1140

Db 1802 TGAGACCCGACGCGCCCTGTCTCAGCTCGTGGTGGCGTTCCGGCGGAGCGGAGCGCGC 1861

QY 1141 CCTTCCCGCAGCCCGCGCGCGCTTCTGCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 1200

Db 1862 CCTTCCCGCAGCCCGCGCGCGCTTCTGCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 1921

QY 1201 CTTCTGCAGTGCCTACGTGCAGCCCTTCTTGGAAGAGCGGCGCTGGAGAGTATC 1260

Db 1922 CTTCTGCAGTGCCTACGTGCAGCCCTTCTTGGAAGAGCGGCGCTGGAGAGTATC 1981

QY 1261 TGTACCCGCGCGGCTGCGCCCGCTTCCGCTGTCTATACCCCGCATCCCGCGCGCGG 1320

Db 1982 TGTACCCGCGCGGCTGCGCCCGCTTCCGCTGTCTATACCCCGCATCCCGCGCGCGG 2041

QY 1321 CGGCAGCCGCGGCGAGCCCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCT 1380

Db 2042 CGGCAGCCGCGGCGAGCCCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCT 2101

QY 1381 CCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440

Db 2102 CCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2161

QY 1441 ACGAGTGGCGCCCTTGGGGCGCGCACCCCGCACCCCGCACCCCGCACCCCGCACCTGC 1500

Db 2162 ACGAGTGGCGCCCTTGGGGCGCGCACCCCGCACCCCGCACCCCGCACCCCGCACCTGC 2221

QY 1501 CCTTCGCGGCGCGCGGAGCGCGGAGCGCGGAGCTCTGCTCAGGAAGATCCCTCGC 1560

Db 2222 CCTTCGCGGCGCGCGGAGCGCGGAGCGCGGAGCTCTGCTCAGGAAGATCCCTCGC 2281

QY 1561 AGCCAGGAAGGAGCTCCCTGA 1583

Db 2282 AGCCAGGAAAGGAAGCTCCCTGA 2304

RESULT 7

ACD19338

ID ACD19338 standard; cDNA; 1471 BP.

XX

AC ACD19338;

XX

DT 25-AUG-2003 (first entry)

XX

DE cDNA encoding novel human protein #18.

XX

KW Human; NOV; gene therapy; endocrine related disease; diabetes;

KW metabolism-related disease; obesity; central nervous system disorder;

KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;

KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;

KW psoriasis; allergy; lupus erythematosus; asthma; cancer;

KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;

KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;

KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;

KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;

KW stroke; infection; gene; ss.

XX

OS Homo sapiens.

XX

PN WO2003023002-A2.

XX

PD 20-MAR-2003.

XX

PF 09-SEP-2002; 2002WO-US028539.

XX

PR 07-SEP-2001; 2001US-0318120P.

PR 07-SEP-2001; 2001US-0318130P.

PR 10-SEP-2001; 2001US-0318430P.

PR 17-SEP-2001; 2001US-0322636P.

PR 17-SEP-2001; 2001US-0322781P.

PR 17-SEP-2001; 2001US-0322816P.

PR 17-SEP-2001; 2001US-0322817P.

PR 19-SEP-2001; 2001US-0323519P.

PR 20-SEP-2001; 2001US-0323631P.

PR 20-SEP-2001; 2001US-0323636P.

PR 25-SEP-2001; 2001US-0324969P.

PR 25-SEP-2001; 2001US-0325091P.

PR 26-SEP-2001; 2001US-0324990P.

PR 17-APR-2002; 2002US-0373212P.

PR 06-SEP-2002; 2002US-00236177.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;

PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;

PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;

PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;

PI Lepley DM, Edinger SR, Burgess CE;

XX

DR WPI; 2003-313242/30.

DR P-PSDB; ABO14645.

XX

PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)

PT and polynucleotides, useful in gene therapy, e.g. for treating or

PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,

PT stroke or infections.

XX

PS Claim 20; Page 125; 586pp; English.

XX

CC The invention describes a new isolated polypeptide (NOVX). The NOVX

CC polypeptide, nucleic acid and antibody are useful as therapeutics,

CC particularly in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, which includes a pathology associated

CC with NOVX polypeptide. The DNA encoding the protein is useful in gene

CC therapy for treating the disease or condition. In particular, the NOVX

CC polypeptide or polynucleotide is useful for treating endocrine/

CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This sequence encodes a novel human NOV protein
XX
SQ Sequence 1471 BP; 288 A; 515 C; 436 G; 232 T; 0 U; 0 Other;
Query Match 40.4%; Score 1471; DB 10; Length 1471;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 129 TTGAACATGGACGAAGAAATTCCTCATTTTGCAGAGAGACAGTTACTGGAAACATAGAGAT 188
Db 1 TTGAACATGGACGAAGAAATTCCTCATTTTGCAGAGAGACAGTTACTGGAAACATAGAGAT 60
QY 189 TTTATAGACTGGACTATTCTCTCTTTGTATATGTGTAAACCCAAAGGAGATGAAACGA 248
Db 61 TTTATAGACTGGACTATTCTCTCTTTGTATATGTGTAAACCCAAAGGAGATGAAACGA 120
QY 249 GACGACACCAAGGATACCTACAAATTACCGACAGATTAAATAGAAAAGAAAGAGAGAC 308
Db 121 GACGACACCAAGGATACCTACAAATTACCGACAGATTAAATAGAAAAGAAAGAGAGAC 180
QY 309 CGAATTAATGAATGCTAGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACA 368
Db 181 CGAATTAATGAATGCTAGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACA 240
QY 369 ACTCTGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACCTTAAAGCT 428
Db 241 ACTCTGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACCTTAAAGCT 300
QY 429 TTAACCGCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGGA 488
Db 301 TTAACCGCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGGA 360
QY 489 TCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGATTTCAAACATGC 548
Db 361 TCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGATTTCAAACATGC 420
QY 549 GCCAAGAGAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGCG 608
Db 421 GCCAAGAGAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGCG 480
QY 609 TGTGTCCAGCTGATCAACCACTTGGACGCGCTGGCCACCCAGTTCTTGCCCAACCCGCGAG 668
Db 481 TGTGTCCAGCTGATCAACCACTTGGACGCGCTGGCCACCCAGTTCTTGCCCAACCCGCGAG 540
QY 669 CTGTTGACTCAACAGGTCCCTCTGAGCAAAAGGACCGGCGCTCCCTCGGCGCGGGGTCC 728
Db 541 CTGTTGACTCAACAGGTCCCTCTGAGCAAAAGGACCGGCGCTCCCTCGGCGCGGGGTCC 600
QY 729 GCGGCGCGCCCTGCTGGAGCGCGGGGGCAGAAAGTGGAGCCCCCTCGCCTACTGCGTG 788
Db 601 GCGGCGCGCCCTGCTGGAGCGCGGGGGCAGAAAGTGGAGCCCCCTCGCCTACTGCGTG 660
QY 789 CCCGTATCCAGCGGACTCAGCCAGCGCCGAGCTCGCGCGAGAACGACACGACACACC 848
Db 661 CCCGTATCCAGCGGACTCAGCCAGCGCCGAGCTCGCGCGAGAACGACACGACACACC 720
QY 849 GACAGCGGCTACGGCGGCGGAAGCCGAGGCCCCGGCCGACCGCGAGAAAGGCAAGCGCG 908
Db 721 GACAGCGGCTACGGCGGCGGAAGCCGAGGCCCCGGCCGACCGCGAGAAAGGCAAGCGCG 780
QY 909 GGGGCGAGCGCGTCAACCATCAAGCAGGAGCCTCCCGGGGAGGACTCGCCGCGGCCAAG 968

Db 781 GGGGAGCGCGCTCACCATCAAGCAGGAGCCTCCCGGGAGGACTCGCGCGGCCCAAG 840
QY 969 AGGATGAAGCTGGATTCCCGCGCGGGCGGCGAGCGCGCGCGCGCGCGCGCGCGCG 1028
Db 841 AGGATGAAGCTGGATTCCCGCGCGGGCGGCGAGCGCGCGCGCGCGCGCGCGCGCG 900
QY 1029 GCGGCGCAGCCCGCTTCTGGGGCCCCGACCCCTGCGCGCGCGCGCGCGCTGTGAGACCC 1088
Db 901 GCGGCGCAGCCCGCTTCTGGGGCCCCGACCCCTGCGCGCGCGCGCGCGCTGTGAGACCC 960
QY 1089 GACGCGCCCTGCTCAGCTCGCTGGTGGCGTTTCGGGGAGGCGGAGGCGCGCTTCCCG 1148
Db 961 GACGCGCCCTGCTCAGCTCGCTGGTGGCGTTTCGGGGAGGCGGAGGCGCGCTTCCCG 1020
QY 1149 CAGCCCGCGCGCGCGCGCGCGCTTCTGCCTGCCCTTCTGCTTCTCTGCTTCTGCA 1208
Db 1021 CAGCCCGCGCGCGCGCGCGCGCTTCTGCCTGCCCTTCTGCTTCTCTGCTTCTGCA 1080
QY 1209 GCTGCGCCTACGTGACGCTTCTTCTGGACAAGAGCGGCGCTGGAGAAGTATCTGTACCCG 1268
Db 1081 GCTGCGCCTACGTGACGCTTCTTCTGGACAAGAGCGGCGCTGGAGAAGTATCTGTACCCG 1140
QY 1269 GCGGCGGCTGCGCGCGCGCTTCTGCTGCTATACCCCGGCATCCCCCGCGCGGCGAGCC 1328
Db 1141 GCGGCGGCTGCGCGCGCGCTTCTGCTGCTATACCCCGGCATCCCCCGCGCGGCGAGCC 1200
QY 1329 GCGGCGCGCGCGCGCGCTTCTGCGCGCGCGCGCGCGCGCTTCTGCTTCTGCTTCTGCGTG 1388
Db 1201 GCGGCGCGCGCGCGCGCTTCTGCGCGCGCGCGCGCGCGCTTCTGCTTCTGCTTCTGCGTG 1260
QY 1389 TTGTGCGCCCTTCCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTGCTTCTGCTTCTGCGTG 1448
Db 1261 TTGTGCGCCCTTCCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTGCTTCTGCTTCTGCGTG 1320
QY 1449 GCGGCTTCTGCTTCTGCGTG 1508
Db 1321 GCGGCTTCTGCTTCTGCGTG 1380
QY 1509 GGGCGCGCGAGCGCGGGGAGCGCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGA 1568
Db 1381 GGGCGCGCGAGCGCGGGGAGCGCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGA 1440
QY 1569 AAGGAAGCTCCCTGAATCCTTGGCTCCCGAA 1599
Db 1441 AAGGAAGCTCCCTGAATCCTTGGCTCCCGAA 1471

RESULT 8
AAF74776
ID AAF74776 standard; cDNA; 1511 BP.
XX
AC AAF74776;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human DEC2b encoding cDNA sequence SEQ ID NO:11.
XX
KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 2..1456
FT /*tag= a
FT /product= "DEC2b"
FT /note= "bHLH type transcription factor"
XX
PN WO200114551-A1.
XX
PD 01-MAR-2001.
XX

PF 19-JUN-2000; 2000WO-JP003991.
XX
PR 19-AUG-1999; 99JP-00233286.
XX
XX (CHUS) CHUGAI SEIYAKU KK.
XX
XX Fujimoto K, Shin M, Kato Y;
XX
XX WPI; 2001-202935/20.
DR P-PSDB; AAB70693.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation.
XX
PS Claim 1; Page 63-68; 83pp; Japanese.
XX
CC The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in the
CC development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC encodes the specifically claimed human DEC2b protein, as given in the
XX present invention
XX
SQ Sequence 1511 BP; 302 A; 521 C; 449 G; 239 T; 0 U; 0 Other;

Query Match 37.8%; Score 1378; DB 4; Length 1511;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAATGAA 320
Db |||||
134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGACCGAATTAATGAA 193

QY 321 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGGACAT 380
Db |||||
194 TGCAATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGGACAT 253

QY 381 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTAACCGCCTTA 440
Db |||||
254 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTAACCGCCTTA 313

QY 441 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATGGGGAGCGGATCTCTGAAATCG 500
Db |||||
314 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATGGGGAGCGGATCTCTGAAATCG 373

QY 501 CCCATTAGTCCGACTTGGATGGCTTCCACTCGGGATTTCAAACATGCGCCAAAGAGTC 560
Db |||||
374 CCCATTAGTCCGACTTGGATGGCTTCCACTCGGGATTTCAAACATGCGCCAAAGAGTC 433

QY 561 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCCGGTGTGTCCAGCTG 620
Db |||||
434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCCGGTGTGTCCAGCTG 493

QY 621 ATCAACCACTTGACGCGCGTGGCCACCCAGTTCTTGGCCACCCCGCAGCTGTGACTCAA 680
Db |||||
494 ATCAACCACTTGACGCGCGTGGCCACCCAGTTCTTGGCCACCCCGCAGCTGTGACTCAA 553

QY 681 CAGTCCCTCTGAGCAAAGGACCGGCGCTCCCTCGGCGCGCGGTCCGCGCGCGCCCGCC 740
Db |||||
554 CAGTCCCTCTGAGCAAAGGACCGGCGCTCCCTCGGCGCGCGGTCCGCGCGCGCCCGCC 613

QY 741 TGCCTGGAGCGCGGGCGGAGAGCTGGAGCCCTCGGCTACTGCGTCCCGTCACTCCAG 800
Db |||||
614 TGCCTGGAGCGCGGGCGGAGAGCTGGAGCCCTCGGCTACTGCGTCCCGTCACTCCAG 673

QY 801 CGGACTCAGCCAGCGCGCGGAGCTCGCGCGCGGAGAACGACACCGACAGCGGCTAC 860
Db |||||
674 CGGACTCAGCCAGCGCGCGGAGCTCGCGCGCGGAGAACGACACCGACAGCGGCTAC 733

QY 861 GCGGCGGAAGCCGAGGCCCGGCGGAGCCGAGAAAGGAGGCGCGGGGCGAGCCGC 920
Db |||||
734 GCGGCGGAAGCCGAGGCCCGGCGGAGCCGCGAGAAAGGAGGCGCGGGGCGAGCCGC 793

QY 921 GTCACCATCAAGCAGAGAGCTCCCGGGAGAGTCTGCCGGCGCCCAAGAGATGAAGCTG 980
Db |||||
794 GTCACCATCAAGCAGAGAGCTCCCGGGAGAGTCTGCCGGCGCCCAAGAGATGAAGCTG 853

QY 981 GATTCCCGCGCGCGGCGGAGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGGCGG 1040
Db |||||
854 GATTCCCGCGCGCGGCGGAGCGCGGCGCGGCGCGGCGCGGCGCGGCGGCGGCGGCGG 913

QY 1041 GCGCTTCTGGGCGCGGACCTCTGCCGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGG 1100
Db |||||
914 GCGCTTCTGGGCGCGGACCTCTGCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 973

QY 1101 CTCAGCTCGCTGGTGGCTTCCGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1160
Db |||||
974 CTCAGCTCGCTGGTGGCTTCCGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1033

QY 1161 GCCGCGGCGCGGCTTCTGCTGCCCTTCTGCTTCTCTCGCTTCTCTCGCTTCTGCGCTAC 1220
Db |||||
1034 GCCGCGGCGCGGCTTCTGCTGCCCTTCTGCTTCTCTCGCTTCTCTCGCTTCTGCGCTAC 1093

QY 1221 GTGCAGCCCTTCTTGACAAAGAGCGGCGGCTGGAGAGTATCTGTATCCCGGCGGCGGCTGCC 1280
Db |||||
1094 GTGCAGCCCTTCTTGACAAAGAGCGGCGGCTGGAGAGTATCTGTATCCCGGCGGCGGCTGCC 1153

QY 1281 GCCCGGCTTCCGCTGCTATACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1340
Db |||||
1154 GCCCGGCTTCCGCTGCTATACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1213

QY 1341 GCCCGGCTTCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1400
Db |||||
1214 GCCCGGCTTCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1273

QY 1401 CCCGAGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1460
Db |||||
1274 CCCGAGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1333

QY 1461 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1520
Db |||||
1334 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1393

QY 1521 CCGGGGAAACCGGAGAGCTCTGTCTCAGGAAGATCCCTCGCAGCCAGGAAGAGCTCCC 1580
Db |||||
1394 CCGGGGAAACCGGAGAGCTCTGTCTCAGGAAGATCCCTCGCAGCCAGGAAGAGCTCCC 1453

QY 1581 TGAATCCTTGGCTCCCGAAGGACGAGGTTCAAGCAGAGTGAGAAAGTTAAATACCCCT 1638
Db |||||
1454 TGAATCCTTGGCTCCCGAAGGACGAGGTTCAAGCAGAGTGAGAAAGTTAAATACCCCT 1511

RESULT 9
AAL26629
ID AAL26629 standard; cDNA; 753 BP.
XX
AC AAL26629;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 19086.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.

PR	29-MAR-2000; 2000US-0193480P.	Db	662 ATACATCCAATTATGATGTCTGACATGC 690
PR	15-MAY-2000; 2000US-0205230P.	RESULT 10	
PR	09-JUN-2000; 2000US-0211315P.	ABX74439	
PR	25-JUL-2000; 2000US-0220534P.	ID	ABX74439 standard; cDNA; 909 BP.
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	XX	ABX74439;
PA	Lillie J, Xu Y, Wang Y, Steinmann K;	AC	ABX74439;
XX	WPI; 2001-451856/48.	XX	21-MAR-2003 (first entry)
XX	New peptide useful as a marker for the diagnosis of breast cancer.	XX	Human cDNA sequence #11 up-regulated in CC-RCC patients.
PT	Claim 1; Page 3602; 3695pp; English.	DE	Human; microarray; solid surface; immobilised probe; CC-RCC;
XX	The invention relates to human breast cancer expressed polynucleotides	XX	differential expression profile; aggressive CC-RCC tumour type;
XX	(AAL07544-AAL26789) and methods of assessing whether a patient is	KW	non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
CC	afflicted with breast cancer by examining the correlation between the	KW	gene expression profiling; tumour tissue; gene; ss.
CC	expression of certain markers and the cancerous state of breast cells.	XX	Homo sapiens.
CC	The polynucleotides and encoded polypeptides are potential markers for	OS	WO200279411-A2.
CC	detecting, diagnosing, monitoring, characterising treating and	XX	10-OCT-2002.
CC	potentially preventing breast cancer. The polynucleotides and encoded	PD	29-MAR-2002; 2002WO-US009576.
CC	polypeptides are also useful for isolating compounds with cytostatic	XX	29-MAR-2001; 2001US-0279411P.
XX	activity	XX	(VAND-) VAN ANDEL INST.
SQ	Sequence 753 BP; 236 A; 151 C; 138 G; 216 T; 0 U; 12 Other;	PI	Haab B, Rhodes D, Teh BT, Takashi M;
	Query Match 17.3%; Score 629; DB 4; Length 753;	XX	WPI; 2003-040679/03.
	Best Local Similarity 100.0%; Pred. No. 8.4e-223;	DR	New microarray, comprising a matrix of cDNA probe from a set of probes
	Matches 629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	immobilized to a solid surface in predetermined order, useful in the
QY	2445 GATTTCTATTTATCCATAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCC 2504	PT	prognosis of patients with clear cell renal carcinoma.
Db	62 GATTTCTATTTATCCATAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCC 121	XX	Claim 35; Page 119-120; 179pp; English.
QY	2505 ACTCTCATTAATCCATCATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGAC 2564	XX	The present invention relates to a microarray comprising a matrix of at
Db	122 ACTCTCATTAATCCATCATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGAC 181	CC	least one cDNA probe from a set of probes immobilised to a solid surface
QY	2565 TGAACCTCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2624	CC	in a predetermined order, where a row of pixels corresponds to replicates
Db	182 TGAACCTCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 241	CC	of one distinct probe from the set. The probes are complementary to
QY	2625 CTATATTTCTGTTAAATTCAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTGTATTGCA 2684	CC	nucleic acid sequences that are expressed differentially in aggressive as
Db	242 CTATATTTCTGTTAAATTCAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTGTATTGCA 301	CC	compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)
QY	2685 AATGCTGTAAAGATATAGTAAGTCTCTGTGAGGCTTCTCTATCTCCAGTCTATGTAT 2744	CC	and that hybridise to the probes under high stringency conditions. The
Db	302 AATGCTGTAAAGATATAGTAAGTCTCTGTGAGGCTTCTCTATCTCCAGTCTATGTAT 361	CC	microarray is useful for the prognosis of patients with CC-RCC, wherein
QY	2745 TTTCTGGAGACCAACCCAGATACCAGATATATCAACAAAGAAAGCTTTTAAATAGGCTTA 2804	CC	aggressive and non-aggressive CC-RCC tumour types are characterised by
Db	362 TTTCTGGAGACCAACCCAGATACCAGATATATCAACAAAGAAAGCTTTTAAATAGGCTTA 421	CC	differential expression profiles of genes that hybridise with one or more
QY	2805 AACCAAGACCTTGTCTAGATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2864	CC	probes immobilised on the microarray. The arrays are useful for gene
Db	422 AACCAAGACCTTGTCTAGATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 481	CC	expression profiling of tumour and normal tissues. The present sequence
QY	2865 ACTTGGATGTTATGTAAGGGGTGAGACACAAAGTCTGACTATGAGTGAGGAAATATCT 2924	XX	represents a human cDNA sequence up-regulated in CC-RCC patients
Db	482 ACTTGGATGTTATGTAAGGGGTGAGACACAAAGTCTGACTATGAGTGAGGAAATATCT 541	SQ	Sequence 909 BP; 218 A; 292 C; 256 G; 143 T; 0 U; 0 Other;
QY	2925 GGGTCTTTTCGTCAGTTTGGTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2984		Query Match 15.2%; Score 552; DB 10; Length 909;
Db	542 GGGTCTTTTCGTCAGTTTGGTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601		Best Local Similarity 100.0%; Pred. No. 2.5e-194;
QY	2985 TGTGTTTAAACAAACGTTAACTCTTAGCCTACAAAGTGGCTCTTATGATAGTATTGTTA 3044		Matches 552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	602 TGTGTTTAAACAAACGTTAACTCTTAGCCTACAAAGTGGCTCTTATGATAGTATTGTTA 661	QY	369 ACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCT 428
QY	3045 ATACATCCAATTAATGATGTCGACATGC 3073	Db	181 ACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCT 240

Db 361 GCCAAGAAGCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCGCGG 420
QY 609 TGTGTCCAGCTGATCAACCACTTGCACGCGGTGGCCACCCAGTTCTTGCCCCACCCCGCAG 668
Db 421 TGTGTCCAGCTGATCAACCACTTGCACGCGGTGGCCACCCAGTTCTTGCCCCACCCCGCAG 480
QY 669 CTGTTGACTCAACAGGTCCCTCTGTAGCRAAAGGCACCGGCGCTCCCTCGGCCCGCGGTCC 728
Db 481 CTGTTGACTCAACAGGTCCCTCTGTAGCRAAAGGCACCGGCGCTCCCTCGGCCCGCGGTCC 540
QY 729 GCGGCCGCCCTCGCTGGAGCGCGGGGCGCAGAAGCTGGAGCCCTCGCCTACTGCGTG 788
Db 541 GCGGCCGCCCTCGCTGGAGCGCGGGGCGCAGAAGCTGGAGCCCTCGCCTACTGCGTG 600
QY 789 CCGGTCAATCCAGCGGACTCAGCCCGAGCGCGGCGCAGAAGCTGGAGCCCTCGCCTACTGCGTG 848
Db 601 CCGGTCAATCCAGCGGACTCAGCCCGAGCGCGGCGCAGAAGCTGGAGCCCTCGCCTACTGCGTG 660
QY 849 GACAGCGGTACGCGCGCGGAAGCCGAGCGCGCGCGCGAGCCGCGGAGAAAGGCGCG 908
Db 661 GACAGCGGTACGCGCGCGGAAGCCGAGCGCGCGCGAGCCGCGGAGAAAGGCGCG 720
QY 909 GGGCGGAGCGCG 920
Db 721 GGGCGGAGCGCG 732

RESULT 11
ADQ24375
ID ADQ24375 standard; DNA; 1573 BP.
XX
AC ADQ24375;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7195.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human; ds.
KW
XX Homo sapiens.
OS
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX

PT Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.
PT
XX
PS Example 2; SEQ ID NO 7195; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic

CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated DNA of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 1573 BP; 508 A; 312 C; 313 G; 411 T; 0 U; 29 Other;
Query Match 14.2%; Score 516; DB 12; Length 1573;
Best Local Similarity 99.8%; Pred. No. 4.9e-181;
Matches 636; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2758 AACGAGATACCAGATAATCACAAAGAAAGCTTTTAAATAGGCTTAAACCAAGACCTTG 2817
Db 251 AACGAGATACCAGATAATCACAAAGAAAGCTTTTAAATAGGCTTAAACCAAGACCTTG 310
QY 2818 TCTAGATATTTTGTAGTTTGTGCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTAT 2877
Db 311 TCTAGATATTTTGTAGTTTGTGCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTAT 370
QY 2878 GTAAGGGGTGAGACACAAACAGTCTGACTATGAGTGAGGAAATATCTGGGCTCTTTTCGTC 2937
Db 371 GTAAGGGGTGAGACACAAACAGTCTGACTATGAGTGAGGAAATATCTGGGCTCTTTTCGTC 430
QY 2938 AGTTTGGTGCAATTTGCTGCTGCTGTTGCTACTGTTTGCCTCAACGCTGTGTTAAACAA 2997
Db 431 AGTTTGGTGCAATTTGCTGCTGCTGTTGCTACTGTTTGCCTCAACGCTGTGTTAAACAA 490
QY 2998 CGTTAAACCTTTAGCCTACAAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTA 3057
Db 491 CGTTAAACCTTTAGCCTACAAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTA 550
QY 3058 ATGATGCTGACATGCTATTTTGTAGGAGAGAAATATGTGCTAATGATATTTTGAGTTA 3117
Db 551 ATGATGCTGACATGCTATTTTGTAGGAGAGAAATATGTGCTAATGATATTTTGAGTTA 610
QY 3118 AAATATCTTTTGGGAGGATTGCTGAAAAGTTGCACCTTTTGTACAATGCTTATGCTTG 3177
Db 611 AAATATCTTTTGGGAGGATTGCTGAAAAGTTGCACCTTTTGTACAATGCTTATGCTTG 670
QY 3178 GTACAAGCTTATGCTGCTTAAATTTTAAATAATTTAAATACTGTCTGTGAGAAACC 3237
Db 671 GTACAAGCTTATGCTGCTTAAATTTTAAATAATTTAAATACTGTCTGTGAGAAACC 729
QY 3238 AGCTGGTTTAAAGAAAGTTTAGTATGTGACGATAAACTAGAAATACCTTATATTCTAGT 3297
Db 730 AGCTGGTTTAAAGAAAGTTTAGTATGTGACGATAAACTAGAAATACCTTATATTCTAGT 789
QY 3298 ATTTTCAGCACTCCATAAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAA 3357
Db 790 ATTTTCAGCACTCCATAAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAA 849
QY 3358 AATTCTTACTAAGGAATAAAAACTTTAATATACGATA 3394
Db 850 AATTCTTACTAAGGAATAAAAACTTTAATATACGATA 886

RESULT 12
AAS26175
ID AAS26175 standard; cDNA; 498 BP.
XX
AC AAS26175;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 354.

XX Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; KW antibacterial; virucide; fungicide; ophthalmological; vulnerary; KW secreted protein; rheumatoid arthritis; hyperproliferative disorder; KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder; KW cerebral ischaemia; angiogenesis; nervous system disorder; KW Alzheimer's disease; infection; ocular disorder; corneal infection;

KW	wound healing; epithelial cell proliferation; skin ageing; food additive;	PR	25-SEP-2000; 2000US-0234997P.
KW	preservative; antiproliferative.	PR	25-SEP-2000; 2000US-0234998P.
XX		PR	26-SEP-2000; 2000US-0235484P.
OS	Homo sapiens.	PR	27-SEP-2000; 2000US-0235834P.
XX		PR	27-SEP-2000; 2000US-0235836P.
PN	WO200155322-A2.	PR	29-SEP-2000; 2000US-0236327P.
XX		PR	29-SEP-2000; 2000US-0236367P.
PD	02-AUG-2001.	PR	29-SEP-2000; 2000US-0236368P.
XX		PR	29-SEP-2000; 2000US-0236369P.
PF	17-JAN-2001; 2001WO-US001341.	PR	29-SEP-2000; 2000US-0236370P.
XX		PR	02-OCT-2000; 2000US-0236802P.
PR	31-JAN-2000; 2000US-0179065P.	PR	02-OCT-2000; 2000US-0237037P.
PR	04-FEB-2000; 2000US-0180628P.	PR	02-OCT-2000; 2000US-0237038P.
PR	24-FEB-2000; 2000US-0184664P.	PR	02-OCT-2000; 2000US-0237039P.
PR	02-MAR-2000; 2000US-0186350P.	PR	02-OCT-2000; 2000US-0237040P.
PR	16-MAR-2000; 2000US-0189874P.	PR	13-OCT-2000; 2000US-0239935P.
PR	17-MAR-2000; 2000US-0190076P.	PR	13-OCT-2000; 2000US-0239937P.
PR	18-APR-2000; 2000US-0198123P.	PR	20-OCT-2000; 2000US-0240960P.
PR	19-MAY-2000; 2000US-0205515P.	PR	20-OCT-2000; 2000US-0241221P.
PR	07-JUN-2000; 2000US-0209467P.	PR	20-OCT-2000; 2000US-0241785P.
PR	28-JUN-2000; 2000US-0214886P.	PR	20-OCT-2000; 2000US-0241786P.
PR	30-JUN-2000; 2000US-0215135P.	PR	20-OCT-2000; 2000US-0241787P.
PR	07-JUL-2000; 2000US-0216647P.	PR	20-OCT-2000; 2000US-0241808P.
PR	07-JUL-2000; 2000US-0216880P.	PR	20-OCT-2000; 2000US-0241809P.
PR	11-JUL-2000; 2000US-0217487P.	PR	20-OCT-2000; 2000US-0241826P.
PR	11-JUL-2000; 2000US-0217496P.	PR	01-NOV-2000; 2000US-0244617P.
PR	14-JUL-2000; 2000US-0218290P.	PR	08-NOV-2000; 2000US-0246474P.
PR	26-JUL-2000; 2000US-0220963P.	PR	08-NOV-2000; 2000US-0246475P.
PR	26-JUL-2000; 2000US-0220964P.	PR	08-NOV-2000; 2000US-0246476P.
PR	14-AUG-2000; 2000US-0224518P.	PR	08-NOV-2000; 2000US-0246477P.
PR	14-AUG-2000; 2000US-0224519P.	PR	08-NOV-2000; 2000US-0246478P.
PR	14-AUG-2000; 2000US-0225213P.	PR	08-NOV-2000; 2000US-0246523P.
PR	14-AUG-2000; 2000US-0225214P.	PR	08-NOV-2000; 2000US-0246524P.
PR	14-AUG-2000; 2000US-0225266P.	PR	08-NOV-2000; 2000US-0246525P.
PR	14-AUG-2000; 2000US-0225267P.	PR	08-NOV-2000; 2000US-0246526P.
PR	14-AUG-2000; 2000US-0225268P.	PR	08-NOV-2000; 2000US-0246527P.
PR	14-AUG-2000; 2000US-0225270P.	PR	08-NOV-2000; 2000US-0246528P.
PR	14-AUG-2000; 2000US-0225447P.	PR	08-NOV-2000; 2000US-0246532P.
PR	14-AUG-2000; 2000US-0225757P.	PR	08-NOV-2000; 2000US-0246609P.
PR	14-AUG-2000; 2000US-0225758P.	PR	08-NOV-2000; 2000US-0246610P.
PR	14-AUG-2000; 2000US-0225759P.	PR	08-NOV-2000; 2000US-0246611P.
PR	18-AUG-2000; 2000US-0226279P.	PR	08-NOV-2000; 2000US-0246613P.
PR	22-AUG-2000; 2000US-0226681P.	PR	17-NOV-2000; 2000US-0249207P.
PR	22-AUG-2000; 2000US-0226688P.	PR	17-NOV-2000; 2000US-0249208P.
PR	22-AUG-2000; 2000US-0227182P.	PR	17-NOV-2000; 2000US-0249209P.
PR	23-AUG-2000; 2000US-0227009P.	PR	17-NOV-2000; 2000US-0249210P.
PR	30-AUG-2000; 2000US-0228924P.	PR	17-NOV-2000; 2000US-0249211P.
PR	01-SEP-2000; 2000US-0229287P.	PR	17-NOV-2000; 2000US-0249212P.
PR	01-SEP-2000; 2000US-0229343P.	PR	17-NOV-2000; 2000US-0249213P.
PR	01-SEP-2000; 2000US-0229344P.	PR	17-NOV-2000; 2000US-0249214P.
PR	01-SEP-2000; 2000US-0229345P.	PR	17-NOV-2000; 2000US-0249215P.
PR	05-SEP-2000; 2000US-0229509P.	PR	17-NOV-2000; 2000US-0249216P.
PR	05-SEP-2000; 2000US-0229513P.	PR	17-NOV-2000; 2000US-0249217P.
PR	06-SEP-2000; 2000US-0230437P.	PR	17-NOV-2000; 2000US-0249218P.
PR	06-SEP-2000; 2000US-0230438P.	PR	17-NOV-2000; 2000US-0249244P.
PR	08-SEP-2000; 2000US-0231242P.	PR	17-NOV-2000; 2000US-0249245P.
PR	08-SEP-2000; 2000US-0231243P.	PR	17-NOV-2000; 2000US-0249264P.
PR	08-SEP-2000; 2000US-0231244P.	PR	17-NOV-2000; 2000US-0249265P.
PR	08-SEP-2000; 2000US-0231413P.	PR	17-NOV-2000; 2000US-0249297P.
PR	08-SEP-2000; 2000US-0231414P.	PR	17-NOV-2000; 2000US-0249299P.
PR	08-SEP-2000; 2000US-0232080P.	PR	17-NOV-2000; 2000US-0249300P.
PR	08-SEP-2000; 2000US-0232081P.	PR	01-DEC-2000; 2000US-0250160P.
PR	12-SEP-2000; 2000US-0231968P.	PR	01-DEC-2000; 2000US-0250391P.
PR	14-SEP-2000; 2000US-0232397P.	PR	05-DEC-2000; 2000US-0251030P.
PR	14-SEP-2000; 2000US-0232398P.	PR	05-DEC-2000; 2000US-0251988P.
PR	14-SEP-2000; 2000US-0232399P.	PR	05-DEC-2000; 2000US-0256719P.
PR	14-SEP-2000; 2000US-0232400P.	PR	06-DEC-2000; 2000US-0251479P.
PR	14-SEP-2000; 2000US-0232401P.	PR	08-DEC-2000; 2000US-0251856P.
PR	14-SEP-2000; 2000US-0233063P.	PR	08-DEC-2000; 2000US-0251868P.
PR	14-SEP-2000; 2000US-0233064P.	PR	08-DEC-2000; 2000US-0251869P.
PR	21-SEP-2000; 2000US-0233065P.	PR	08-DEC-2000; 2000US-0251989P.
PR		PR	08-DEC-2000; 2000US-0251990P.
PR		PR	11-DEC-2000; 2000US-0254097P.

PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-488783/53.
DR P-PSDB; AAU16188.
DR
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
PS Claim 1; SEQ ID NO 354; 980pp; English.
PS
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 13.6%; Score 496; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.4e-173;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 TGAACGAGACGACACCAAGGATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAA 300
Db 1 TGAACGAGACGACACCAAGGATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAA 60
QY 301 GAAGAGACCGGAATTAATGAATGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGA 360
Db 61 GAAGAGACCGGAATTAATGAATGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGA 120
QY 361 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGAATTAACTTTGAAACACT 420
Db 121 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGAATTAACTTTGAAACACT 180
QY 421 TAAAAGCTTTAACCGCCCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATG 480
Db 181 TAAAAGCTTTAACCGCCCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATG 240
QY 481 GGGAGCGATCTCTGAAATCGGCCATTTCAGTCCGACTTGGATCGGTTCCACTCGGGATTTT 540
Db 241 GGGAGCGATCTCTGAAATCGGCCATTTCAGTCCGACTTGGATCGGTTCCACTCGGGATTTT 300
QY 541 AAACATGCGCCAAAGAGCTTTCGAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 600
Db 301 AAACATGCGCCAAAGAGCTTTCGAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 360
QY 601 AGCCGCGGTGTGTCCAGCTGATCAACCACTTGCAAGCGGTGGCCACCCAGTTCTTGCCCA 660
Db 361 AGCCGCGGTGTGTCCAGCTGATCAACCACTTGCAAGCGGTGGCCACCCAGTTCTTGCCCA 420
QY 661 CCCCGAGCTGTTGACTCAACAGGTTCCCTCTGAGCAAAAGGCACCGGCTCCCTCGGCCG 720

Db 421 CCCCGAGAGCTGTTGACTCAACAGGTCCTCTGAGCAAAAGSCACCGGCTCCCTCGGCCG 480
QY 721 CCGGGTCCGGCGGCCG 736
Db 481 CCGGGTCCGGCGGCCG 496
RESULT 13
ABX73516
ID ABX73516 standard; DNA; 498 BP.
XX
AC ABX73516;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polynucleotide #344.
XX
KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PN US2002132753-A1.
XX
PD 19-SEP-2002.
XX
PF 17-JAN-2001; 2001US-00764864.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-147444/14.
DR P-PSDB; ABU55256.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 354; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 498 BP; 142 A; 137 C; 115 G; 104 T; 0 U; 0 Other;

Query Match 13.6%; Score 496; DB 8; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.4e-173;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 TGAACGAGACGACCAAGGATACCTACAATTACCGCACAGATTATAGAAAAGAAA 300
Db 1 TGAACGAGACGACCAAGGATACCTACAATTACCGCACAGATTATAGAAAAGAAA 60

Qy 301 GAAGAGACCGAATTAATGAATGCATTGCTCAGCTGAAAGATTTACTGCCTGAACATCTGA 360
Db 61 GAAGAGACCGAATTAATGAATGCATTGCTCAGCTGAAAGATTTACTGCCTGAACATCTGA 120

Qy 361 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGAATTAACTTTGAAACACT 420
Db 121 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGAATTAACTTTGAAACACT 180

Qy 421 TAAAAGCTTTAACCGCCTTAACCGAGCAACAGCATCAGAAGATAAATTCGCTTACAGAATG 480
Db 181 TAAAAGCTTTAACCGCCTTAACCGAGCAACAGCATCAGAAGATAAATTCGCTTACAGAATG 240

Qy 481 GGGAGCGATCTCTGAATCGCCCATTCAGTCCGACTTGGATGCGTTCACACTCGGGATTTC 540
Db 241 GGGAGCGATCTCTGAATCGCCCATTCAGTCCGACTTGGATGCGTTCACACTCGGGATTTC 300

Qy 541 AAACATGCGCCAAAGAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 600
Db 301 AAACATGCGCCAAAGAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 360

Qy 601 AGCCGGGTGTTCAGCTGATCAACCACTTGCACGCCGTGGCCACCCAGTTCTTGCCCA 660
Db 361 AGCCGGGTGTTCAGCTGATCAACCACTTGCACGCCGTGGCCACCCAGTTCTTGCCCA 420

Qy 661 CCCCAGCTGTTGACTCAACAGGTCCTCTGAGCAAAGGCACCGGCGCTCCCTCGGCCG 720
Db 421 CCCCAGCTGTTGACTCAACAGGTCCTCTGAGCAAAGGCACCGGCGCTCCCTCGGCCG 480

Qy 721 CCGGGTCCGCGCGCGC 736
Db 481 CCGGGTCCGCGCGCGC 496

RESULT 14
ACC72415
ID ACC72415 standard; cDNA; 430 BP.
XX
AC ACC72415;
XX
DT 07-JUL-2003 (first entry)
XX
DE Human cDNA insert from ovarian cancer clone O1668P.
XX
KW Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer; gene;
KW ss.
XX
OS Homo sapiens.
XX
PN WO2003029468-A1.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031467.
XX
PR 02-OCT-2001; 2001US-0327135P.
PR 30-MAY-2002; 2002US-0384531P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Mannion J;
XX
DR WPI; 2003-372001/35.
XX
PT New polynucleotide and polypeptide useful for diagnosing and/or treating
PT cancer, particularly ovarian cancer, and as a vaccine.
XX
PS Claim 1; Page 147; 169pp; English.
XX
CC The invention relates to a novel isolated polynucleotide. The
CC polynucleotides of the invention have cytostatic activity, and may have a
CC use in gene therapy, and in a vaccine. The composition and methods are
CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.
CC The composition may also be used as a vaccine to prevent cancer. The
CC present sequence is used in the exemplification of the invention
XX
SQ Sequence 430 BP; 162 A; 81 C; 61 G; 126 T; 0 U; 0 Other;

Query Match 11.8%; Score 430; DB 9; Length 430;
Best Local Similarity 100.0%; Pred. No. 3.9e-149;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2240 AAATGGTATCTCTTAGTAACCTTGACCTCGTAAAGAAACACGGAGCTGGCCATCGTCAG 2299
Db 1 AAATGGTATCTCTTAGTAACCTTGACCTCGTAAAGAAACACGGAGCTGGCCATCGTCAG 60

Qy 2300 AACTAAGTCAGGAGAGAGATGGATGAGAGCCAGAAATCATTCCTAGTACATTGCTAA 2359
Db 61 AACTAAGTCAGGAGAGAGATGGATGAGAGCCAGAAATCATTCCTAGTACATTGCTAA 120

Qy 2360 CACTTTATTGAGAAATTGACCATGAATTAATGGAATCATCTTAATTTCTTAAGTCCAT 2419
Db 121 CACTTTATTGAGAAATTGACCATGAATTAATGGAATCATCTTAATTTCTTAAGTCCAT 180

Job time : 1668.37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 17, 2004, 03:15:04 ; Search time 15113.1 Seconds
(without alignments)
11392.870 Million cell updates/sec

Title: US-10-078-650-1
Perfect score: 3641
Sequence: 1 ctgcactgaaggaggagc.....ttaaaaaaaaaaaaaaa 3641

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_cm:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3641	100.0	3641	6	BD011858	BD011858 Novel bHL
2	3641	100.0	3641	9	AB044088	AB044088 Homo sapi
3	3043	83.6	3431	9	BC025968	BC025968 Homo sapi
4	2853	78.4	204228	9	AC022509	AC022509 Homo sapi
5	2290	62.9	3274	6	BD176928	BD176928 Novel bHL
6	1583	43.5	2304	6	AX591648	AX591648 Sequence
7	1378	37.8	1511	6	BD011867	BD011867 Novel bHL
8	921	25.3	1973	9	AK094422	AK094422 Homo sapi
9	629	17.3	753	6	CQ434066	CQ434066 Sequence
10	552	15.2	909	6	CQ727066	CQ727066 Sequence
11	460	12.6	593	9	HSA270695	AJ270695 Homo sapi
12	413	11.3	797	11	BV208394	BV208394 BHLHB3_20
13	391	10.7	462	6	CQ424141	CQ424141 Sequence
14	386	10.6	447	6	CQ432989	CQ432989 Sequence
15	363	10.0	471	6	CQ685084	CQ685084 Sequence
16	355	9.8	521	6	AX198436	AX198436 Sequence
17	354	9.7	501	6	AX208985	AX208985 Sequence
18	354	9.7	501	6	AX209655	AX209655 Sequence
19	341	9.4	580	6	AX198458	AX198458 Sequence

20	341	9.4	580	6	AX209005	AX209005 Sequence
21	339	9.3	411	6	CQ415230	CQ415230 Sequence
22	312	8.6	497	6	CQ680194	CQ680194 Sequence
C 23	308	8.5	553	6	CQ410788	CQ410788 Sequence
C 24	308	8.5	684	6	CQ398115	CQ398115 Sequence
C 25	308	8.5	684	6	CQ404404	CQ404404 Sequence
26	296	8.1	409	6	AX209663	AX209663 Sequence
27	296	8.1	502	6	AX198354	AX198354 Sequence
28	296	8.1	502	6	AX208910	AX208910 Sequence
29	281	7.7	400	6	CQ712682	CQ712682 Sequence
30	274	7.5	285	6	AX209128	AX209128 Sequence
31	274	7.5	397	6	AX198590	AX198590 Sequence
C 32	271	7.4	322	6	AX208788	AX208788 Sequence
C 33	271	7.4	424	6	AX198219	AX198219 Sequence
C 34	266	7.3	440	6	AX341473	AX341473 Sequence
35	265	7.3	404	6	CQ682847	CQ682847 Sequence
36	252	6.9	278	6	CQ695845	CQ695845 Sequence
37	246	6.8	502	6	AX892348	AX892348 Sequence
38	246	6.8	502	6	BD027881	BD027881 Sequence
39	242	6.6	604	6	CQ412716	CQ412716 Sequence
40	240	6.6	364	6	CQ697079	CQ697079 Sequence
41	239	6.6	480	6	CQ685093	CQ685093 Sequence
42	237	6.5	261	6	CQ687963	CQ687963 Sequence
43	234	6.4	260	6	CQ695847	CQ695847 Sequence
44	226	6.2	260	6	CQ688754	CQ688754 Sequence
45	225	6.2	265	6	CQ686173	CQ686173 Sequence

ALIGNMENTS

RESULT 1
BD011858
LOCUS BD011858 3641 bp DNA linear PAT 02-AUG-2002
DEFINITION Novel bHLH type transcriptional gene, DEC2.
ACCESSION BD011858
VERSION BD011858.1 GI:22092047
KEYWORDS WO 0114551-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3641)
AUTHORS Fujimoto,K., Shin,M. and Kato,Y.
TITLE Novel bHLH type transcriptional gene, DEC2
JOURNAL Patent: WO 0114551-A 1 01-MAR-2001;
CHUGAI PHARM CO LTD,KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
COMMENT OS Homo sapiens (human)
PN WO 0114551-A/1
PD 01-MAR-2001
PF 19-JUN-2000 WO 2000JP003991
PR 19-AUG-1999 JP 99P 233286
PI KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
PC C12N15/12,C07L14/475,C07K16/18
CC
FH
FT

FEATURES

source
Location/Qualifiers
1..3641
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 3641; DB 6; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCACTGAAGAGGAGCGGAGAGAGAGAGAGACTGGAGACGCACAGATCCCCCAAGGTCTC 60

Db 1 CTGCACTGAAGAGGAGCGGAGAGAGAGAGACTGGAGACGCACAGATCCCCCAAGGTCTC 60

QY 61 CCAAGCCTACCGTCCACAGATTATTGTACAGAGCCCCCAAAATCGAAACAGAGGAAACG 120

Db	61	CCAAGCTACCGTCCCAACAGATTATTGTACAGAGCCCAAAAATCGAAACAGAGGAAACG	120
QY	121	AACAGCAGTTGAACATGGACGAAGGAATTCCCTCATTTGCAAGAGAGACAGTACTGGAAC	180
Db	121	AACAGCAGTTGAACATGGACGAAGGAATTCCCTCATTTGCAAGAGAGACAGTACTGGAAC	180
QY	181	ATAGAGATTTTATAGGACTGGACTATTCCCTCTTTGTATATGTGTAAACCCAAAAGGAGCA	240
Db	181	ATAGAGATTTTATAGGACTGGACTATTCCCTCTTTGTATATGTGTAAACCCAAAAGGAGCA	240
QY	241	TGAACGAGACGACACCAAGGATACCTACAAATTTACCGCACAGATTAAATAGAAAAGAAA	300
Db	241	TGAACGAGACGACACCAAGGATACCTACAAATTTACCGCACAGATTAAATAGAAAAGAAA	300
QY	301	GAAGAGACCGAATTAATGAATGCATTGCTCAGCTGAAAGATTTACTGCCTGAACATCTGA	360
Db	301	GAAGAGACCGAATTAATGAATGCATTGCTCAGCTGAAAGATTTACTGCCTGAACATCTGA	360
QY	361	AATTGACAACTCTGGGACATCTGGAGAAAAGCTGTAGTCTTGGAAATTAACITTTGAAACACT	420
Db	361	AATTGACAACTCTGGGACATCTGGAGAAAAGCTGTAGTCTTGGAAATTAACITTTGAAACACT	420
QY	421	TAAAAGCTTTAACCGCTTAAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATG	480
Db	421	TAAAAGCTTTAACCGCTTAAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATG	480
QY	481	GGGAGCGATCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGGATTTT	540
Db	481	GGGAGCGATCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGGATTTT	540
QY	541	AAACATGCGCCAAAGAAAGTCTTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGG	600
Db	541	AAACATGCGCCAAAGAAAGTCTTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGG	600
QY	601	AGCCGGGTGTGCCAGTCAACAGGTCCCTCTGAGCAAAAGGCACCGGCTCCCTCGGCCG	660
Db	601	AGCCGGGTGTGCCAGTCAACAGGTCCCTCTGAGCAAAAGGCACCGGCTCCCTCGGCCG	660
QY	661	CCCGGAGCTGTTGACTCAACAGGTCCCTCTGAGCAAAAGGCACCGGCTCCCTCGGCCG	720
Db	661	CCCGGAGCTGTTGACTCAACAGGTCCCTCTGAGCAAAAGGCACCGGCTCCCTCGGCCG	720
QY	721	CCGGTCCGCGCGCCCTGCTGGAGCGCGCGGCGCAGAGCTGGAGCCCTCGCCT	780
Db	721	CCGGTCCGCGCGCCCTGCTGGAGCGCGCGGCGCAGAGCTGGAGCCCTCGCCT	780
QY	781	ACTGGTGCCCGTCACTCCAGCGGACTCAGCCCGAGCGCGGAGCTCGCCCGAGAACGACA	840
Db	781	ACTGGTGCCCGTCACTCCAGCGGACTCAGCCCGAGCGCGGAGCTCGCCCGAGAACGACA	840
QY	841	CGGACCCGACAGCGGCTACGGCGCGGAAGCCGAGGCGCCGCGGACCCGCGAGAAAGGCA	900
Db	841	CGGACCCGACAGCGGCTACGGCGCGGAAGCCGAGGCGCCGCGGACCCGCGAGAAAGGCA	900
QY	901	AAGCGCGGGGGAGCGCGCTACCCATCAAGCAGGAGCCTCCCGGGGAGGACTCGCCGG	960
Db	901	AAGCGCGGGGGAGCGCGCTACCCATCAAGCAGGAGCCTCCCGGGGAGGACTCGCCGG	960
QY	961	CGCCAAAGAGGATGAAGCTGGATTCCCGCGCGCGCAGCGCGCGCGCGGGGGGGG	1020
Db	961	CGCCAAAGAGGATGAAGCTGGATTCCCGCGCGCGCAGCGCGCGCGCGGGGGGGG	1020
QY	1021	GCGCGCGCGCGCGGAGCGCGCTTCTGGGGCCGACCCCTGCCCGCGCGCGCGCTGC	1080
Db	1021	GCGCGCGCGCGCGGAGCGCGCTTCTGGGGCCGACCCCTGCCCGCGCGCGCGCTGC	1080
QY	1081	TGAGACCCGACCGCCCTGCTCAGCTCGCTGGTGGCTTGGCGGAGGCGGCGCGC	1140
Db	1081	TGAGACCCGACCGCCCTGCTCAGCTCGCTGGTGGCTTGGCGGAGGCGGCGCGC	1140
QY	1141	CCTTCCGCGACCGCGCGCGCGCGCGCTTCTGCTTCCCTCTCTCGC	1200

Db	1141	CCTTCCGCGACCGCGCGCGCGCGCGCTTCTGCTTCCCTCTCTCGC	1200
QY	1201	CTTCTGCAGTCCCGCTACGTGCAGCCCTTCTCTGACAAGAGCGGCTTGAGAAATATC	1260
Db	1201	CTTCTGCAGTCCCGCTACGTGCAGCCCTTCTCTGACAAGAGCGGCTTGAGAAATATC	1260
QY	1261	TGTACCCGCGCGGCTGCCGCCCTTCCCGCTGCTATACCCCGGCATCCCGCCCCCGG	1320
Db	1261	TGTACCCGCGCGGCTGCCGCCCTTCCCGCTGCTATACCCCGGCATCCCGCCCCCGG	1320
QY	1321	CGGAGCCGCGGACGCGCGCGCGCTGCGCGCGCGCGCGCGCTTCCCTGCCTGT	1380
Db	1321	CGGAGCCGCGGACGCGCGCGCGCTGCGCGCGCGCGCGCGCTTCCCTGCCTGT	1380
QY	1381	CCTCGGTGTTGTGCGCCCTCCGAGAAGGGCGCGCGCGCGCGCTTCTCTCTGCGCG	1440
Db	1381	CCTCGGTGTTGTGCGCCCTCCGAGAAGGGCGCGCGCGCGCGCTTCTCTCTGCGCG	1440
QY	1441	ACGAGGTGGCGCCCTTGGGGCGCGCACCCCGACCGCGCGCGCGCGCGCTTCTCTCTG	1500
Db	1441	ACGAGGTGGCGCCCTTGGGGCGCGCACCCCGACCGCGCGCGCGCGCTTCTCTCTG	1500
QY	1501	CCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTCTCTGCGCG	1560
Db	1501	CCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTCTCTGCGCG	1560
QY	1561	AGCAGGAAAGGAAGTCCCTGAATCCTTGCCTCCGAGGAGCGGAGGTTCAAGCAGAGT	1620
Db	1561	AGCAGGAAAGGAAGTCCCTGAATCCTTGCCTCCGAGGAGCGGAGGTTCAAGCAGAGT	1620
QY	1621	GAGAAGTTAAATACCTTAAGGAGGTTCAAGCAGAGTGAGAAGTTAAATACCTTAAG	1680
Db	1621	GAGAAGTTAAATACCTTAAGGAGGTTCAAGCAGAGTGAGAAGTTAAATACCTTAAG	1680
QY	1681	GTCTTTAAGGGAGGAAGTGTATAGATGCACGACAGGCATAAACAAGAACAAACAG	1740
Db	1681	GTCTTTAAGGGAGGAAGTGTATAGATGCACGACAGGCATAAACAAGAACAAACAG	1740
QY	1741	GTGTATGTGTACATTCGGAGTTCCTGTTTGTCTCATCCCGCACCCCGCTCCACA	1800
Db	1741	GTGTATGTGTACATTCGGAGTTCCTGTTTGTCTCATCCCGCACCCCGCTCCACA	1800
QY	1801	CCTAACATCCCTTCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCTTCTCTCTCTCT	1860
Db	1801	CCTAACATCCCTTCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCTTCTCTCTCTCT	1860
QY	1861	TCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1920
Db	1861	TCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1920
QY	1921	TAAATTTGAAACCTAATTCGAGAGGAAGTAAAGAGGTTAGGCGATGTTCCATATCTCT	1980
Db	1921	TAAATTTGAAACCTAATTCGAGAGGAAGTAAAGAGGTTAGGCGATGTTCCATATCTCT	1980
QY	1981	GAAACCCCGGGTAGGGGAAAGATGTTAAACCTTTGACGCTTTTGGAGTTGACATGGAAC	2040
Db	1981	GAAACCCCGGGTAGGGGAAAGATGTTAAACCTTTGACGCTTTTGGAGTTGACATGGAAC	2040
QY	2041	AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAGTGCGCTGCTTCTGAGGCGCTAGGT	2100
Db	2041	AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAGTGCGCTGCTTCTGAGGCGCTAGGT	2100
QY	2101	TCCACAAACAGATTGAGGCTCTTTTAGAATTTACTCTTCTCAGTATTTCTTAATGT	2160
Db	2101	TCCACAAACAGATTGAGGCTCTTTTAGAATTTACTCTTCTCAGTATTTCTTAATGT	2160
QY	2161	TCAGCTTCTTAAAGGCATATATTTTCAAAGAGTGAGGATGCGAGTTTCTCAGGTTGCA	2220
Db	2161	TCAGCTTCTTAAAGGCATATATTTTCAAAGAGTGAGGATGCGAGTTTCTCAGGTTGCA	2220
QY	2221	ACCTATTCTGAAGTGTAAATGGTATCTCTTAGTAACTTGCACTCGTTAAAGAAACAC	2280
Db	2221	ACCTATTCTGAAGTGTAAATGGTATCTCTTAGTAACTTGCACTCGTTAAAGAAACAC	2280

QY 2281 GGAGCTGGGCCATCGT CAGAACTAAGTCAGGGAAGGAGATGGATGAGAAAGGCCAGAAATCA 2340
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2281 GGAGCTGGGCCATCGT CAGAACTAAGTCAGGGAAGGAGATGGATGAGAAAGGCCAGAAATCA 2340
QY 2341 TTCCTAGTACATTTGCTAACACCTTTATTTGAGAAAATTGACCATGAATTAATGGACTCATCT 2400
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2341 TTCCTAGTACATTTGCTAACACCTTTATTTGAGAAAATTGACCATGAATTAATGGACTCATCT 2400
QY 2401 TAATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA 2460
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2401 TAATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA 2460
QY 2461 TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCACTCTCATTAATCCAT 2520
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2461 TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCACTCTCATTAATCCAT 2520
QY 2521 CATGTTCCCTAAATTTTGTAAATCTTACTGTAAAAAAAGTGCACTGAACTTCAAAACAAA 2580
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2521 CATGTTCCCTAAATTTTGTAAATCTTACTGTAAAAAAAGTGCACTGAACTTCAAAACAAA 2580
QY 2581 ACAAACAAACAAACAAACAAACAAACAACTGATATATCTCTATATCTCTGTTAAAA 2640
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2581 ACAAACAAACAAACAAACAAACAAACAACTGATATATCTCTATATCTCTGTTAAAA 2640
QY 2641 TTCAAAAGTGAACGAAAGCATTTTAAGTGGCCAGTTTTTGAATGTCGAAATGCTGTAAGATAT 2700
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2641 TTCAAAAGTGAACGAAAGCATTTTAAGTGGCCAGTTTTTGAATGTCGAAATGCTGTAAGATAT 2700
QY 2701 AGAATGAAGTCTCTGTGAGGCCCTTCTATCTCCAAGTCTATGTATTTTCTGGAGACCAAC 2760
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2701 AGAATGAAGTCTCTGTGAGGCCCTTCTATCTCCAAGTCTATGTATTTTCTGGAGACCAAC 2760
QY 2761 CAGATACAGATAATCACAAGAAAGCTTTTTTAATAAGGCTAAACCAAGACCTTGCT 2820
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2761 CAGATACAGATAATCACAAGAAAGCTTTTTTAATAAGGCTAAACCAAGACCTTGCT 2820
QY 2821 AGATATTTTGTAGTTTGTGCCAAGTGTGAGAAATCTCACTGGATGTTATGTA 2880
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2821 AGATATTTTGTAGTTTGTGCCAAGTGTGAGAAATCTCACTGGATGTTATGTA 2880
QY 2881 AGGGGTGAGACACAACAGTCTGACTATGAGTGAGGAAAATATCTGGTCTTTTCGTCAGT 2940
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2881 AGGGGTGAGACACAACAGTCTGACTATGAGTGAGGAAAATATCTGGTCTTTTCGTCAGT 2940
QY 2941 TTGGTGCAATTTGCTGCTGCTGTTGTGCTACTGTTTGCCTCAACGCTGTGTTTAAACAACGT 3000
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2941 TTGGTGCAATTTGCTGCTGCTGTTGTGCTACTGTTTGCCTCAACGCTGTGTTTAAACAACGT 3000
QY 3001 TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAAATTAATG 3060
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3001 TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAAATTAATG 3060
QY 3061 ATGCTGACATGCTATTTTGTAGGAGAAAATATGTGCTAATGATATTTTGAGTTAAAA 3120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3061 ATGCTGACATGCTATTTTGTAGGAGAAAATATGTGCTAATGATATTTTGAGTTAAAA 3120
QY 3121 TATCTTTTGGGAGGATTTGCTGAAAAGTTGCACCTTTTGTACAAATGCTTATGCTTGGTA 3180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3121 TATCTTTTGGGAGGATTTGCTGAAAAGTTGCACCTTTTGTACAAATGCTTATGCTTGGTA 3180
QY 3181 CAAGCCTATGCTCTTAAATTTATTTAAAAAAATTAATAACTGTCTGTGAGAAACACAGC 3240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3181 CAAGCCTATGCTCTTAAATTTATTTAAAAAAATTAATAACTGTCTGTGAGAAACACAGC 3240
QY 3241 TGGTTTAAAAAGTTTAGTATGTGACGATAAACTAGAAATACCTTTATATTTCTAGTATT 3300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3241 TGGTTTAAAAAGTTTAGTATGTGACGATAAACTAGAAATACCTTTATATTTCTAGTATT 3300
QY 3301 TTCAGCACTCCATAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT 3360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3301 TTCAGCACTCCATAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT 3360

QY 3361 TCTTACTAAGGATAAAAACTTTTAATATACGATATGATATTGTCTAATAATTAATAAAGA 3420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3361 TCTTACTAAGGATAAAAACTTTTAATATACGATATGATATTGTCTAATAATTAATAAAGA 3420
QY 3421 CATAATGGATGCTCAATTAGTTTTTAAGATATCTATAACTATAGGGATACAAATCACTACA 3480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3421 CATAATGGATGCTCAATTAGTTTTTAAGATATCTATAACTATAGGGATACAAATCACTACA 3480
QY 3481 GTTCTCAGATTTACACCTTTTTTTTTGTCATTTGGCTTGATGTCACACATTTCCAATCTCTT 3540
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3481 GTTCTCAGATTTACACCTTTTTTTTTGTCATTTGGCTTGATGTCACACATTTCCAATCTCTT 3540
QY 3541 GCAAGCCTCCAGCTCTGGCTTTTGTCTACCTGCTCGTTCCCAATGTATCTTAATGAAAAG 3600
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3541 GCAAGCCTCCAGCTCTGGCTTTTGTCTACCTGCTCGTTCCCAATGTATCTTAATGAAAAG 3600
QY 3601 TGCAAAAGAAAACCTACCAATTAATAAAAAAATAAAAAA 3641
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3601 TGCAAAAGAAAACCTACCAATTAATAAAAAAATAAAAAA 3641

RESULT 2
AB044088
LOCUS AB044088 Homo sapiens mRNA for bHLH protein DEC2, complete cds.
DEFINITION Homo sapiens mRNA for bHLH protein DEC2, complete cds.
ACCESSION AB044088
VERSION AB044088.1 GI:12583668
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Fujimoto,K., Shen,M., Noshiro,M., Matsubara,K., Shingu,S.,
Honda,K., Yoshida,E., Suardita,K., Matsuda,Y. and Kato,Y.
TITLE Molecular cloning and characterization of DEC2, a new member of
basic helix-loop-helix proteins
JOURNAL Biochem. Biophys. Res. Commun. 280 (1), 164-171 (2001)
MEDLINE 21092582
PUBMED 11162494
REFERENCE 2 (bases 1 to 3641)
AUTHORS Fujimoto,K.
TITLE Direct Submission
JOURNAL Submitted (31-MAY-2000) Katsumi Fujimoto, hiroshima university,
department of biochemistry, school of dentistry; 1-2-3, kasumi,
minami-ku, hiroshima 734-8553, Japan
(E-mail:kfujimo@hiroshima-u.ac.jp, Tel:81822575629,
Fax:81822575629)

FEATURES
source Location/Qualifiers
1. .3641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1. .3641
/gene="DEC2"
CDS 135. .1583
/gene="DEC2"
/codon_start=1
/product="bHLH protein DEC2"
/protein_id="BAB21502.1"
/db_xref="GI:12583669"
translation="MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDITKDT
YKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHKLALTAL
TEOQHQKIIALQNGERSLKSPIQSDLDADFHSGFQICAKEVLQYLSRFESWTPREPCV
QLINLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAPCLERAGQKLEPLAYCV
PVIQRTQPSAELEAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPGEDSPA
PKMKLDSRGGGGGGGAAAAAALLGPPDPAALALRPDAALLSSLVAFGGGG
APFPQAAAAAPFLCFLSPSAAAAAYVQFLDKSGLEKLYLPAAAAAAPFLLYPGI
PAPAAAAAATAAAAAAFAFCLSSVLSPPPEKAGAAAATLLPHEVAPLGAHPQHPH
GRTHLPFAGPREPGNPESQAQEDPSQPGKEAP"

ORIGIN
Query Match 100.0%; Score 3641; DB 9; Length 3641;

Best Local Similarity 100.0%; Pred. No. 0; Matches 3641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	CTGCACTGAAGAGGGAGACGCGAGAGAGACTTGAGACGCACAGATCCCCCAAGTCTC	60	
Db	1	CTGCACTGAAGAGGGAGACGCGAGAGAGACTTGAGACGCACAGATCCCCCAAGTCTC	60	
QY	61	CCAAGCCTACCCTCCACAGATTATTGTACAGAGCCCCAAAATCGAAACAGAGGAAACG	120	
Db	61	CCAAGCCTACCCTCCACAGATTATTGTACAGAGCCCCAAAATCGAAACAGAGGAAACG	120	
QY	121	AACAGCAGTTGAAACATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAC	180	
Db	121	AACAGCAGTTGAAACATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAC	180	
QY	181	ATAGAGATTTTATAGACTGGACTATTCTCTTTGTATATGTGTAAACCCAAAAGGAGCA	240	
Db	181	ATAGAGATTTTATAGACTGGACTATTCTCTTTGTATATGTGTAAACCCAAAAGGAGCA	240	
QY	241	TGAAACGAGACGACACCAAGGATACCTACAAATTAACCGCACAGATTATAGAAAAGAAA	300	
Db	241	TGAAACGAGACGACACCAAGGATACCTACAAATTAACCGCACAGATTATAGAAAAGAAA	300	
QY	301	GAAGAGACCGAATTAATGAATGCATTGCTCAGCTCAAAGATTACTGCCTGAACATCTGA	360	
Db	301	GAAGAGACCGAATTAATGAATGCATTGCTCAGCTCAAAGATTACTGCCTGAACATCTGA	360	
QY	361	AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAATTAACTTTGAACACT	420	
Db	361	AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAATTAACTTTGAACACT	420	
QY	421	TAAAAGCTTTAACCGCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATG	480	
Db	421	TAAAAGCTTTAACCGCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATG	480	
QY	481	GGGAGCGATCTCTGAATTCGCCCATTCAGTCCGACTTGGATCGTTCCACTCGSGATTTC	540	
Db	481	GGGAGCGATCTCTGAATTCGCCCATTCAGTCCGACTTGGATCGTTCCACTCGSGATTTC	540	
QY	541	AAACATGCGCCAAAGAGTCTTGCAATACCTCTCCCCTTTGAGAGCTGGACACCCAGGG	600	
Db	541	AAACATGCGCCAAAGAGTCTTGCAATACCTCTCCCCTTTGAGAGCTGGACACCCAGGG	600	
QY	601	AGCCGCGTGTGTCAGTGATCAACCACTTGCAACCGCTGGCCACCCAGTTCTTGCCCA	660	
Db	601	AGCCGCGTGTGTCAGTGATCAACCACTTGCAACCGCTGGCCACCCAGTTCTTGCCCA	660	
QY	661	CCCGCAGCTGTTGACTCAACAGGTCCCTCTGAGCAAGGCACCGCGCTCCCTCGGCGG	720	
Db	661	CCCGCAGCTGTTGACTCAACAGGTCCCTCTGAGCAAGGCACCGCGCTCCCTCGGCGG	720	
QY	721	CCGGTCCGCGCGCCOCTGCTGGAGCGCGGGCAGAAAGCTGGAGCCCCTCGCCT	780	
Db	721	CCGGTCCGCGCGCCOCTGCTGGAGCGCGGGCAGAAAGCTGGAGCCCCTCGCCT	780	
QY	781	ACTGCGTCCCCTCATCCAGCGGACTCAGCCCAGCGCGAGCTCGCCCGAGAACGACA	840	
Db	781	ACTGCGTCCCCTCATCCAGCGGACTCAGCCCAGCGCGAGCTCGCCCGAGAACGACA	840	
QY	841	CGGACACCGACAGCGGCTACCGCGGCGGAAGCCGAGGCCCGGCCAGAACGSCA	900	
Db	841	CGGACACCGACAGCGGCTACCGCGGCGGAAGCCGAGGCCCGGCCAGAACGSCA	900	
QY	901	AAGGCGGGGGCGAGCCCGTCAACCATCAAGCAGGAGSCTCCCGGGAGGACTCGCGGG	960	
Db	901	AAGGCGGGGGCGAGCCCGTCAACCATCAAGCAGGAGSCTCCCGGGAGGACTCGCGGG	960	
QY	961	CGCCCAAGAGGATGAAGCTGGATTCCCGCGGCGGCGGAGCGGCGGCCGGGGCGG	1020	
Db	961	CGCCCAAGAGGATGAAGCTGGATTCCCGCGGCGGCGGAGCGGCGGCCGGGGCGG	1020	
QY	1021	GCGCGGCGGCGGAGCCGCGCTTCTGGGGCCCGACCTGCCGCCCGCGCGCGCTGC	1080	

QY	2161	TCAGCTTTCTAAAGGCATATATTTTTTCAAAGAAGTGAGGATGCAGTTTCTCAGCTTGCA	2220
Db	2161	TCAGCTTTCTAAAGGCATATATTTTTTCAAAGAAGTGAGGATGCAGTTTCTCAGCTTGCA	2220
QY	2221	ACCTATTCTGAAGTGGTTTTAAATGGTAICTCTTAGTAACCTTGCACCTCGTTAAAGAAACAC	2280
Db	2221	ACCTATTCTGAAGTGGTTTTAAATGGTAICTCTTAGTAACCTTGCACCTCGTTAAAGAAACAC	2280
QY	2281	GGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAAATCA	2340
Db	2281	GGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAAATCA	2340
QY	2341	TTCTTAGTACATTGCTTAACACTTTTATGAGAAATTTGACCATGAATTAATGGACTCATCT	2400
Db	2341	TTCTTAGTACATTGCTTAACACTTTTATGAGAAATTTGACCATGAATTAATGGACTCATCT	2400
QY	2401	TAATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA	2460
Db	2401	TAATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA	2460
QY	2461	TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCACCTCTCATTAATCCAT	2520
Db	2461	TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCACCTCTCATTAATCCAT	2520
QY	2521	CATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGCACCTGAACTTTCAAAACAAA	2580
Db	2521	CATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGCACCTGAACTTTCAAAACAAA	2580
QY	2581	ACAAAAACAACAACAACAACAAAGTCCAAACTGATATATCTCTATATTTCTGTTAAAA	2640
Db	2581	ACAAAAACAACAACAACAACAAAGTCCAAACTGATATATCTCTATATTTCTGTTAAAA	2640
QY	2641	TTCAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTGTATTGCAAAATGCTGTAAAGATAT	2700
Db	2641	TTCAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTGTATTGCAAAATGCTGTAAAGATAT	2700
QY	2701	AGAATGAAGTCTGTGAGGCCCTTCCTATCTCCAAGTCTATGTATTTTCTGGAGACCAAC	2760
Db	2701	AGAATGAAGTCTGTGAGGCCCTTCCTATCTCCAAGTCTATGTATTTTCTGGAGACCAAC	2760
QY	2761	CAGATACCAGATAATCACAAAGAAAGCTTTTAAATAGGCTTAAACCAAGACCTTGCT	2820
Db	2761	CAGATACCAGATAATCACAAAGAAAGCTTTTAAATAGGCTTAAACCAAGACCTTGCT	2820
QY	2821	AGATATTTTACTTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTA	2880
Db	2821	AGATATTTTACTTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTA	2880
QY	2881	AGGGGTGAGACACAACAGTCTGACTATGAGTGAGGAAAAATATCTGGTCTTTTCGTCAGT	2940
Db	2881	AGGGGTGAGACACAACAGTCTGACTATGAGTGAGGAAAAATATCTGGTCTTTTCGTCAGT	2940
QY	2941	TTGGTGCAATTGCTGCTGCTGTTGTGCTACTGTTTGCCTCAAACGCTGTGTTTAAACACGT	3000
Db	2941	TTGGTGCAATTGCTGCTGCTGTTGTGCTACTGTTTGCCTCAAACGCTGTGTTTAAACACGT	3000
QY	3001	TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG	3060
Db	3001	TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG	3060
QY	3061	ATGCTGCACATGCTATTTTGTAGGGAGAAAAATATGTCTAATGATATTTTGAGTTAAAA	3120
Db	3061	ATGCTGCACATGCTATTTTGTAGGGAGAAAAATATGTCTAATGATATTTTGAGTTAAAA	3120
QY	3121	TATCTTTTGGGAGGATTTGCTGAAAAGTGCACCTTTTGTACAAATGCTTATGCTTGTA	3180
Db	3121	TATCTTTTGGGAGGATTTGCTGAAAAGTGCACCTTTTGTACAAATGCTTATGCTTGTA	3180
QY	3181	CAAGCTTATGCTGCTTAAATTTTAAAAAATTAATAATCTGCTGTGAGAAACACGAGC	3240
Db	3181	CAAGCTTATGCTGCTTAAATTTTAAAAAATTAATAATCTGCTGTGAGAAACACGAGC	3240

QY	3241	TGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAATTACCTTTATATTCTAGTATT	3300
Db	3241	TGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAATTACCTTTATATTCTAGTATT	3300
QY	3301	TTCAGCACTCCATAAATTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT	3360
Db	3301	TTCAGCACTCCATAAATTCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAAT	3360
QY	3361	TCTTACTAAGGAATAAAAACTTTTAATATACGATATGATATTGTCTAATAATTAATAAAGA	3420
Db	3361	TCTTACTAAGGAATAAAAACTTTTAATATACGATATGATATTGTCTAATAATTAATAAAGA	3420
QY	3421	CATAATGGATGCTCAATTAGTTTAAAGATATCTATAACTATAGGGATACAAATCACTACA	3480
Db	3421	CATAATGGATGCTCAATTAGTTTAAAGATATCTATAACTATAGGGATACAAATCACTACA	3480
QY	3481	GTTCTCAGATTTACACCTTTTTTTTGTCTATGGCTTGATGTCACACATTTCCATCTCTT	3540
Db	3481	GTTCTCAGATTTACACCTTTTTTTTGTCTATGGCTTGATGTCACACATTTCCATCTCTT	3540
QY	3541	GCAAGCCTCCAGGCTCTGGCTTTTGTCTACCTGCTGTTCCCAATGTATCTTAATGAAAAG	3600
Db	3541	GCAAGCCTCCAGGCTCTGGCTTTTGTCTACCTGCTGTTCCCAATGTATCTTAATGAAAAG	3600
QY	3601	TGCAAAAGAAAAACCTACCAATTAAAAAATAAAAAA 3641	
Db	3601	TGCAAAAGAAAAACCTACCAATTAAAAAATAAAAAA 3641	

RESULT 3	
BC025968	3431 bp mRNA linear PRI 30-JUN-2004
LOCUS	
DEFINITION	Homo sapiens basic helix-loop-helix domain containing, class B, 3, mRNA (CDNA clone MGC:39365 IMAGE:4650325), complete cds.
ACCESSION	BC025968
VERSION	BC025968.1 GI:19684063
KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 3431) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
AUTHORS	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 3431) Strausberg,R.
AUTHORS	Direct Submission
TITLE	Submitted (18-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
JOURNAL	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk

Email: ccapbs-r@mail.nih.gov

Tissue Procurement: ATCC

Tissue Procurement: ATCC

cdNA Library Preparation: Rubin Laboratory

cdNA Library Arrayed by: The I.M.A.G.E. Co.

DNA Sequencing by: Genome Sequence Cent

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
Parvaneh Saeedi, JR Santos, Angélique Schnerch, Ursula Skalska,
Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
Schnein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 42 Row: f Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.

FEATURES

source

```
1. .3431
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:39365 IMAGE:4650325"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH MGC 21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1. .3431
```

gene

```
1. .3431
/gene="BHLHB3"
/note="synonyms: SHARP-1, DEC2, SHARP1"
/db_xref="LocusID:79365"
/db_xref="MIM:606200"
```

53

```

292. .1740
/gene="BHLHB3"
/codon_start=1
/product="basic helix-loop-helix domain containing, class
B, 3"
/protein_id="AAH25968.1"
/db_xref="GI:19684064"
/db_xref="LocusID:79365"
/db_xref="MIM:606200"
/translation="MDEGIPLHQERQLLEHRDFIGLDYSSLYMCKPKRSMKRD DTKD
YKLPRLHLEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL
TEQHQKI IALQNGSSSPKI QSDLDFAFGSGFTCAKEVLQYLSRFESWTPREPRC
QLINHLHAVA TQFLPTQLLTQQVPLSGTGA SGAASAAAPCLERAQKLEPLAYC
PVIOTQPSAE LAEAENDITDSDSGYGEAEARPDREKKGAGASRVTIKQEPFGEDSP
PKENKLD SRGGSGGGPGGGA VAAAALLGPD PAAAAAALLRPDAALLSSLVAFGGGG
APFPQAAAAA AAFCLPFCFLSPSAAAA YVQPFLLDKSGLEKLYLPAAAAAPFLLYPG
PAPAAAAA AAAAAA AAFCLPSSVSPPPKAGAAA AATLLPHEVAPLGA PHPQHP
GRTHLPFAGPREPGNPESSAOEDSPGPGKEAP"

```

ORIGIN

	Query Match	83.6%;	Score 3043;	DB 9;	Length 3431;
	Best Local Similarity	99.9%;	Pred. No. 0;		
	Matches 3213;	Conservative	0;	Mismatches	1; Indels 1; Gaps 1;
QY	1	CTGCAC	TGAAGGGAGCGGAGAGAGAGACTGGAGACGCACAGATCCCCC	CAAGGTCTC	60
DB	158	CTGCAC	TGAAGGGAGCGGAGAGAGAGACTGGAGACGCACAGATCCCCC	CAAGGTCTC	217
QY	61	CCAAGCCT	ACCGTCCCACAGATTATTGTACAGAGCCCCCAAAATCGAAACAGAGGAAACG	120	
DB	218	CCAAGCCT	ACCGTCCCACAGATTATTGTACAGAGCCCCCAAAATCGAAACAGAGGAAACG	277	
QY	121	AACAGCAG	TTGAACATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAC	180	
DB	278	AACAGCAG	TTGAACATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAC	337	

Db 1418 TGTACCCGGGGGGCTGCCGCCCGTTCCCGTGTATACCCGGGCATCCCGCGCCCGG 1477
QY 1321 CGGCAGCCGGGGCAGCCCGCCCGCGCTGCCCGCCCGCCCGCGCGTTCCTGCTGT 1380
Db 1478 CGGCAGCCGGGGCAGCCCGCCCGCGCTGCCCGCCCGCGCGCGTTCCTGCTGT 1537
QY 1381 CCTCGGTGTGTGCCCCCTCCCGAGAGGGGGGGCGCGCGCGCGCGACCTCCTGCCGC 1440
Db 1538 CCTCGGTGTGTGCCCCCTCCCGAGAGGGGGGGCGCGCGCGCGCGACCTCCTGCCGC 1597
QY 1441 ACGAGGTGGCGCCCTTGGGGCGCGCGCACCCCGCAGCACCCCGCAGCGCGCGCACCTGCG 1500
Db 1598 ACGAGGTGGCGCCCTTGGGGCGCGCGCACCCCGCAGCACCCCGCAGCGCGCGCACCTGCG 1657
QY 1501 CCTTCGCGGGCCCCCGGAGCGCGGGAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 1560
Db 1658 CCTTCGCGGGCCCCCGGAGCGCGGGAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 1717
QY 1561 AGCCAGGAAGGAAGCTCCCTGAATCCTTTCGTCGCCGAGGAGGTTCAAGCAGAGT 1620
Db 1718 AGCCAGGAAGGAAGCTCCCTGAATCCTTTCGTCGCCGAGGAGGTTCAAGCAGAGT 1777
QY 1621 GAGAAATTAAATACCCCTTAAGGAGGTTCAAGCAGAGTGAGAAGTTAAATACCCCTAAG 1680
Db 1778 GAGAAATTAAATACCCCTTAAGGAGGTTCAAGCAGAGTGAGAAGTTAAATACCCCTAAG 1837
QY 1681 GTCTTTAAGGAGGAAGTGTAATAGATGCACGACAGGCATAAACAAGAAACAACAACAG 1740
Db 1838 GTCTTTAAGGAGGAAGTGTAATAGATGCACGACAGGCATAAACAAGAAACAACAACAG 1897
QY 1741 GTGTTATGTGTACATTCGGAGTTCCTGTTTTCGTCATCCCGCACCCACCCCTCCACA 1800
Db 1898 GTGTTATGTGTACATTCGGAGTTCCTGTTTTCGTCATCCCGCACCCACCCCTCCACA 1957
QY 1801 CACTAACATCCCTTCTTCCCGCCCGCAGCTGTAAAAGATCCTATGCGAAAGACACTGGC 1860
Db 1958 CACTAACATCCCTTCTTCCCGCCCGCAGCTGTAAAAGATCCTATGCGAAAGACACTGGC 2017
QY 1861 TCCTTTTTTAATCCCCCAATAAATTTTGCCCCCTTTTAGGCCATGTTCCATTATCTCT 1920
Db 2018 TCCTTTTTTAATCCCCCAATAAATTTTGCCCCCTTTTAGGCCATGTTCCATTATCTCT 2077
QY 1921 TAAATTTGAAACCTAATTCGAGAGGAAGTAAGAAGGTCGTCTGTGGCTGAGCTAGGT 1980
Db 2078 TAAATTTGAAACCTAATTCGAGAGGAAGTAAGAAGGTCGTCTGTGGCTGAGCTAGGT 2137
QY 1981 GAACCCCGGGTAGGGGAAAGATGTTAAACACCTTTGACGTCCTTTGGAGTTGACATGGAAC 2040
Db 2138 GAACCCCGGGTAGGGGAAAGATGTTAAACACCTTTGACGTCCTTTGGAGTTGACATGGAAC 2197
QY 2041 AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCTGTTTAGGAGCGGT 2100
Db 2198 AGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCTGTTTAGGAGCGGT 2257
QY 2101 TCCACAAACAGATTGAGGCTCTTTTGTAGAATTGAATTTACTCTTCAGTATTTTCTAATGT 2160
Db 2258 TCCACAAACAGATTGAGGCTC-TTTTGAATTTGAATTTACTCTTCAGTATTTTCTAATGT 2316
QY 2161 TCAGCTTTCTAAAAGGCATATATTTTCAAAGAAAGTGAGGATGCAGTTTCTCACGTTGCA 2220
Db 2317 TCAGCTTTCTAAAAGGCATATATTTTCAAAGAAAGTGAGGATGCAGTTTCTCACGTTGCA 2376
QY 2221 ACCTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAACCTTGCACTCGTTAAAGAAACAC 2280
Db 2377 ACCTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAACCTTGCACTCGTTAAAGAAACAC 2436
QY 2281 GGAGCTGGGCCCATCGTCAGAACTTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAAATCA 2340
Db 2437 GGAGCTGGGCCCATCGTCAGAACTTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAAATCA 2496
QY 2341 TTCCCTAGTACATTTGCTAACACCTTTATTGAGAAATTTGACCATGAATTAATGGACTCATCT 2400

Db 2497 TTCTTAGTACATTTGCTAACACACTTTTATTGAGAAATTGACCATGAATTAATGGACTCATCT 2556
QY 2401 TAAATTTCTTAAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA 2460
Db 2557 TAAATTTCTTAAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA 2616
QY 2461 TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCACTCTCATTAATCCAT 2520
Db 2617 TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCACTCTCATTAATCCAT 2676
QY 2521 CATGTTCCCTAAATTTTGTAAATCTTACTGTAAAAAAAAGTGCACCTGAACTTCAAAACAAA 2580
Db 2677 CATGTTCCCTAAATTTTGTAAATCTTACTGTAAAAAAAAGTGCACCTGAACTTCAAAACAAA 2736
QY 2581 ACAAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 2640
Db 2737 ACAAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 2796
QY 2641 TTCAAAAAGTGAACGAAGCATTTTAACTGGCCAGTTTTTGGATTGCAATGCTGTAAAGATAT 2700
Db 2797 TTCAAAAAGTGAACGAAGCATTTTAACTGGCCAGTTTTTGGATTGCAATGCTGTAAAGATAT 2856
QY 2701 AGAATGAAGTCTGTGAGGCCCTTCCCTATCTCCTCAAGTCTATGTATTTTCTGGAGACCAAC 2760
Db 2857 AGAATGAAGTCTGTGAGGCCCTTCCCTATCTCCTCAAGTCTATGTATTTTCTGGAGACCAAC 2916
QY 2761 CAGATACCAGATAATCACAAAGAAAGCTTTTAAATAGGCTTTAAACCAAGACCTTGTCT 2820
Db 2917 CAGATACCAGATAATCACAAAGAAAGCTTTTAAATAGGCTTTAAACCAAGACCTTGTCT 2976
QY 2821 AGATATTTTGTAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTA 2880
Db 2977 AGATATTTTGTAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTA 3036
QY 2881 AGGGGTGAGACAAACAGTCTGACTATGAGTGAGGAAAAATATCTGGGTCTTTTCGTCAGT 2940
Db 3037 AGGGGTGAGACAAACAGTCTGACTATGAGTGAGGAAAAATATCTGGGTCTTTTCGTCAGT 3096
QY 2941 TTGGTGCAATTTGCTGCTGCTGTTGCTACTGTTTGCCCTCAAACGCTGTGTTTAAACCAACGT 3000
Db 3097 TTGGTGCAATTTGCTGCTGCTGTTGCTACTGTTTGCCCTCAAACGCTGTGTTTAAACCAACGT 3156
QY 3001 TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG 3060
Db 3157 TAAACTCTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATG 3216
QY 3061 ATGCTGTACATGCTATTTTGTAGGAGAAAAATATGTGCTAATGATATTTTGAGTTAAAA 3120
Db 3217 ATGCTGTACATGCTATTTTGTAGGAGAAAAATATGTGCTAATGATATTTTGAGTTAAAA 3276
QY 3121 TATCTTTTGGGAGGATTTTGCTGAAAAAGTTGCACCTTTTGTTAACATGCTTATGTTGTA 3180
Db 3277 TATCTTTTGGGAGGATTTTGCTGAAAAAGTTGCACCTTTTGTTAACATGCTTATGTTGTA 3336
QY 3181 CAAGCTTATGCTGCTTAAATTTTAAAAAAAAT 3215
Db 3337 CAAGCTTATGCTGCTTAAATTTTAAAAAAAAT 3371

RESULT 4
AC022509/c
LOCUS
DEFINITION
AC022509
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

204228 bp
Homo sapiens 12 BAC RP11-283G6 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
AC022509
AC022509.21 GI:11465000
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 204228)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulsged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shoostari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 204228)
Worley,K.C.
Direct Submission
Submitted (05-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 204228)
Worley,K.C.
Direct Submission
Submitted (30-NOV-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 30, 2000 this sequence version replaced gi:10945701.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the

EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Option Info -----
Version: 1.01
\$Id: bcm-qual-stats.cc,v 1.6 1998/10/27 00:34:03 kdurbin Exp
kdurbin \$

Options Specified: xfo
XGapFile: ./xgap/hadw.final
OutFile :
FA File : ./xgap/HADW.fa
Sequence start: 10
Sequence end: 204237
Optimistic: 0
Use no qual cols: 0

----- Summary Statistics -----

Contig length: 204228
Phrap values in estimate: 203313
Average error rate (BCM-Phrap estimate): 0.000109636
Fraction of Phrap values less than 40 : 0.0179674
Number of consensus changing edits: 37
Number of N's in consensus : 0
Clip start: 11
Clip end: 204238Num reads in contig :
2905
Num reads missing quality: 174
Num reads invalid quality: 0

----- Low Quality Bases -----

Quality	Position	Surrounding Sequence
4	6263	ttttttttt(t)gagacagggg
7	89744	aattaataac(a)agggcagctg
7	89830	ctggtagttt(g)agaccagcct
7	156088	gacaatcaga(a)atgaacttaa
8	6188	aatcactaag(c)taaaagaaaa
8	7146	ccactatcac(a)agaatagcat
8	89632	gtcaggaaaa(t)gccatctgtt
8	89633	tcaggaatat(g)ccatctgtta
8	89634	caggaaaatg(c)catctgttac
8	89745	attaataaca(a)gggcagctgg
8	89791	cagcattttg(g)gaggtgagg
8	89792	agcattttgg(g)aggtgaggt
8	89795	attttgggag(g)ctgaggtggg
8	89829	tctggtagtt(t)gagaccagcc
8	140166	gggggagggg(g)gaggggggag
8	140167	gggggagggg(g)aggggggagg
9	6160	aaagaaaaa(t)aatcttggg
9	7030	gtggaagtg(a)aggggaagca
9	7097	gagggagcgc(a)aagggagaag

9	7099	gggagcgcaa(a)gggagaagt	
9	7106	caagggaga(a)gtgccacatc	
9	7128	gatctcatga(c)aactccccca	
9	7130	tctcatgaca(a)ctccccact	
9	7135	tgacaactcc(c)ccactatcac	
9	7136	gacaactccc(c)cactatcaca	
9	7137	acaactcccc(c)actatcacaa	
9	7138	caactcccc(a)ctatcacaa	
9	7139	aactccccca(c)tatcacaa	
9	7144	ccccactatc(a)caagaatagc	
9	7148	actatcacaa(g)aatagcatgg	
9	7149	ctatcacaa(g)atagcatggg	
9	7150	tatcacaa(g)a)tagcatggg	
9	89628	ccagtcagg(a)aatgccatc	
9	89768	atggtggctt(a)cagctgtaat	
9	89769	tgggtgctta(c)agctgtaatc	
9	89793	gcattttggg(a)ggctgaggtg	
9	89828	gtctggtagt(t)tgagaccagc	
9	89865	aaccccatct(c)tacaaaaaat	
10	6159	gaaaggaaaa(a)taaatcttgg	
10	6179	ggccccaaa(a)tcactaaagt	
10	6180	ggccccaaaa(t)cactaaagta	
10	6209	aagtgaagct(g)ggaactgctt	
10	6222	aactgcttg(g)gcaactgcc	
10	6234	caactgctt(c)ctatttttt	
10	6540	agccaccgtg(c)ctaccgcct	
10	6542	ccaccgtgcc(c)taccgctcc	
10	7029	ggtggaaggt(g)aaggggaagc	
10	7057	gtctttacaa(g)gtggcaggag	
10	7058	tccttacaag(g)tggcaggagg	
10	7085	gagagagaga(c)agagggaagc	
10	7086	agagagagac(a)gagggaagc	
10	7087	gagagagaca(g)agggagcgca	
10	7096	agaggagcg(c)aaagggaagaa	
10	7100	ggagcgcaaa(g)ggagaagtgc	
10	7101	gagcgcaaa(g)gagaagtgc	
10	7102	agcgcaaa(g)gagaagtgc	
10	7103	gcgcaaaagg(a)gaagtgccac	
10	7104	cgcaaaaggga(g)aagtgccaca	
10	7105	gcaaaaggga(a)agtgccacat	
10	7131	ctcatgacaa(c)tcctccacta	
10	7142	tccccacta(t)cacaagaata	
Query Match 78.4%; Score 2853; DB 9; Length 204228;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 3143; Conservative 0; Mismatches 1; Indels 2; Gaps 2;			
QY	480	GGGGAGCGATCTCTGAAATCGCCCATTCAGTCCGACTTGGAATCGCGATTCAGTCCGATTCGCGATTT	539
Db	44987	GGGGAGCGATCTCTGAAATCGCCCATTCAGTCCGACTTGGAATCGCGATTCAGTCCGATTT	44928
QY	540	CAACATCGGCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGG	599
Db	44927	CAACATCGGCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGG	44868
QY	600	GAGCCGCGTGTGTCCAGCTGATCAACCACTTGACGCCCGTGGCCACCCAGTTCTTGCCC	659
Db	44867	GAGCCGCGTGTGTCCAGCTGATCAACCACTTGACGCCCGTGGCCACCCAGTTCTTGCCC	44808
QY	660	ACCCCGCAGTGTGACTCAACAGTCCCTCTGAGCAAAAGGACCGCGCTCCCTCGGCC	719
Db	44807	ACCCCGCAGTGTGACTCAACAGTCCCTCTGAGCAAAAGGACCGCGCTCCCTCGGCC	44748
QY	720	GCCGGTCCGGCGCGCCCCCTGCTGGAGCGCGGGGACAGAGCTGGAGCCCTCGGC	779
Db	44747	GCCGGTCCGGCGCGCCCCCTGCTGGAGCGCGGGGACAGAGCTGGAGCCCTCGGC	44688
QY	780	TACTGCTGCCGCTCATCCAGCGGACTCAGCCAGCGCCGAGCTCGCCCGGAGACGAC	839
Db	44687	TACTGCTGCCGCTCATCCAGCGGACTCAGCCAGCGCCGAGCTCGCCCGGAGACGAC	44628
QY	840	ACGGACACCGACAGCGGCTACGGCGGCGAAGCCGAGGCCCGCGGACCGCGAGGAAGGC	899

Db	44627	ACGGACACCGACAGCGGTACGGCGGCGAAGCCGAGGCCCGCGGACCGGAGAAAAGGC	44568
QY	900	AAAGGCGCGGGGAGCGCGTCAACCATCAAGCAGGAGCCTCCCGGGAGGACTCGCCG	959
Db	44567	AAAGGCGCGGGGAGCGCGTCAACCATCAAGCAGGAGCCTCCCGGGAGGACTCGCCG	44508
QY	960	GCGCCCAAGAGGATGAAGCTGGATTCCCGCGCGGCGGAGCGCGGCGCGCGGGGC	1019
Db	44507	GCGCCCAAGAGGATGAAGCTGGATTCCCGCGCGGCGGAGCGCGGCGCGGGGC	44448
QY	1020	GCGCGCGCGCGGCGGAGCGCGCTTCTGGGGCCCGACCCCTGCGCGCGCGCGCGCTG	1079
Db	44447	GCGCGCGCGCGGCGGAGCGCGCTTCTGGGGCCCGACCCCTGCGCGCGCGCGCGCTG	44388
QY	1080	CTGAGACCCGACGCGCGCTGCTCAGCTCGCTGGTGGCGTTCCGCGGAGCGGAGCGCG	1139
Db	44387	CTGAGACCCGACGCGCGCTGCTCAGCTCGCTGGTGGCGTTCCGCGGAGCGGAGCGCG	44328
QY	1140	CCCTTCCCGCAGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTGCTTCTCTCTCG	1199
Db	44327	CCCTTCCCGCAGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTGCTTCTCTCTCG	44268
QY	1200	CCTTCTGCAGCTGCGCGCTACGTGCAGCCCTTCTGGACAAGAGCGGCTTGAGAAATAT	1259
Db	44267	CCTTCTGCAGCTGCGCGCTACGTGCAGCCCTTCTGGACAAGAGCGGCTTGAGAAATAT	44208
QY	1260	CTGTACCCGCGCGCGCTGCGCGCGCGCTGCTGCTATACCCCGGATCCCCGCCCG	1319
Db	44207	CTGTACCCGCGCGCGCTGCGCGCGCGCTGCTGCTATACCCCGGATCCCCGCCCG	44148
QY	1320	GCGGACGCGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCTGCTGCTG	1379
Db	44147	GCGGACGCGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCTGCTGCTG	44088
QY	1380	TCCTCGGTGTGTCGCCCTCCCGAGAGCGCGCGCGCGCGCGCGCTGCTGCTGCG	1439
Db	44087	TCCTCGGTGTGTCGCCCTCCCGAGAGCGCGCGCGCGCGCGCGCTGCTGCTGCG	44028
QY	1440	CACGAGGTGGCGCGCTTGGGGCGCGCACCCCGAGCACCCCGAGCGCGCACCCACCTG	1499
Db	44027	CACGAGGTGGCGCGCTTGGGGCGCGCACCCCGAGCACCCCGAGCGCGCACCCACCTG	43968
QY	1500	CCCTTCGCGCGCGCGCGAGCGCGGGAACCCCGAGAGCTCTGCTCAGGAAGATCCCTCG	1559
Db	43967	CCCTTCGCGCGCGCGCGAGCGCGGGAACCCCGAGAGCTCTGCTCAGGAAGATCCCTCG	43908
QY	1560	CAGCCAGGAAAGGAAGCTCCCTGAATCCTTGCCTCCGAGAGCGAGGTTCAAGCAGAG	1619
Db	43907	CAGCCAGGAAAGGAAGCTCCCTGAATCCTTGCCTCCGAGAGCGAGGTTCAAGCAGAG	43848
QY	1620	TGAGAAAGTTAAATACCTTAAAGGAGTTCAAGCAGAGTGAGAGTTAAATACCTTAA	1679
Db	43847	TGAGAAAGTTAAATACCTTAAAGGAGTTCAAGCAGAGTGAGAGTTAAATACCTTAA	43788
QY	1680	GGTCTTTAAGGAGGAAGTGTAAATAGATGCACGACAGGCATAAACAACAACAACA	1739
Db	43787	GGTCTTTAAGGAGGAAGTGTAAATAGATGCACGACAGGCATAAACAACAACAACA	43728
QY	1740	GGTGTATGTGTACATTGGGAGTTCTGTTTGTCTCATCCCGCACCCACCCCTCCAC	1799
Db	43727	GGTGTATGTGTACATTGGGAGTTCTGTTTGTCTCATCCCGCACCCACCCCTCCAC	43668
QY	1800	ACACTAATCCCTTTCTTCCCGCACCCAGCTGTAAAGATCTATGCGAAAGACACTGG	1859
Db	43667	ACACTAATCCCTTTCTTCCCGCACCCAGCTGTAAAGATCTATGCGAAAGACACTGG	43608
QY	1860	CTCTTTTAAATCCCGCAATAAATTTTGCCTTTAGGCAATGTTCCATTATCTC	1919
Db	43607	CTCTTTTAAATCCCGCAATAAATTTTGCCTTTAGGCAATGTTCCATTATCTC	43548
QY	1920	TTAAATTTGGAACCTAATTCGAGAGGAAGTAAGAGGCTCTGTTCTGTGGCTGAGCTAGG	1979
Db	43547	TTAAATTTGGAACCTAATTCGAGAGGAAGTAAGAGGCTCTGTTCTGTGGCTGAGCTAGG	43488

Qy	1980	TGAACCCCGGGGTAGGGGAAAGATGTTAACACCTTTGACGCTCTTTGGAGTTGACATGGAA	2039
Db	43487	TGAACCCCGGGGTAGGGGAAAGATGTTAACACCTTTGACGCTCTTTGGAGTTGACATGGAA	43428
Qy	2040	CAGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCCCTGTTTAGGAGGCG	2099
Db	43427	CAGCAGGTAGTTGTTATGTAGAGCTAGTTCTCAAAGCTGCCCTGCCCTGTTTAGGAGGCG	43368
Qy	2100	TTCCACAAACAGATTGAGGCTCTTTTGTAGAAATTGAATTTACTCTTTCAGTATTTTCTAATG	2159
Db	43367	TTCCACAAACAGATTGAGGCTC-TTTTGAATTTGAATTTACTCTTTCAGTATTTTCTAATG	43309
Qy	2160	TTCAGCTTTCTAAAAGGCATATATTTTCAAAGAGTGAGGATGCAGTTTCTCACGTTGC	2219
Db	43308	TTCAGCTTTCTAAAAGGCATATATTTTCAAAGAGTGAGGATGCAGTTTCTCACGTTGC	43249
Qy	2220	AACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAACA	2279
Db	43248	AACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAACA	43189
Qy	2280	CGGAGCTGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAATC	2339
Db	43188	CGGAGCTGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAATC	43129
Qy	2340	ATTCTAGTACATTTGCTAACACTTTATTGAGAAATTGACCATGAATTAATGGACTCATC	2399
Db	43128	ATTCTAGTACATTTGCTAACACTTTATTGAGAAATTGACCATGAATTAATGGACTCATC	43069
Qy	2400	TTAATTTCTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTATCC	2459
Db	43068	TTAATTTCTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTATCC	43009
Qy	2460	ATAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCACCTCATTAATCCA	2519
Db	43008	ATAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCACCTCATTAATCCA	42949
Qy	2520	TCATGTTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGCACTGAACCTTCAAAACAA	2579
Db	42948	TCATGTTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGCACTGAACCTTCAAAACAA	42889
Qy	2580	AACAAAAACAACAACAACAAAGTCCAACTGATATATCCTATATTTCTGTTAAA	2639
Db	42888	AACAAAAACAACAACAACAAAGTCCAACTGATATATCCTATATTTCTGTTAAA	42829
Qy	2640	ATTCAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTGTATTGCAAAATGCTGTAAAGATA	2699
Db	42828	ATTCAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTGTATTGCAAAATGCTGTAAAGATA	42769
Qy	2700	TAGAAATGAAGTCTGTGAGGCTTCTATCTCCAACTATGATATTTTCTGGAGACCAAA	2759
Db	42768	TAGAAATGAAGTCTGTGAGGCTTCTATCTCCAACTATGATATTTTCTGGAGACCAAA	42709
Qy	2760	CCAGATACCAGATAATCAAAAAGAAAGCTTTTAAAGGCTTAAACCAAGACCTTGTGTC	2819
Db	42708	CCAGATACCAGATAATCAAAAAGAAAGCTTTTAAAGGCTTAAACCAAGACCTTGTGTC	42649
Qy	2820	TAGATATTTTGTGTTGGCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGT	2879
Db	42648	TAGATATTTTGTGTTGGCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGT	42589
Qy	2880	AAGGGGTGAGACACAACAGTCTGACTATGAGTGAGGAAAAATATCTGGGCTTTTTCGTGAG	2939
Db	42588	AAGGGGTGAGACACAACAGTCTGACTATGAGTGAGGAAAAATATCTGGGCTTTTTCGTGAG	42529
Qy	2940	TTTGGTGCATTTGCTGCTGTTGCTACTGTTTGGCTCAAAACGCTGTGTTTAAACAACG	2999
Db	42528	TTTGGTGCATTTGCTGCTGTTGCTACTGTTTGGCTCAAAACGCTGTGTTTAAACAACG	42469
Qy	3000	TTAAACTCTTAGCCTACAAAGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAAT	3059
Db	42468	TTAAACTCTTAGCCTACAAAGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAAT	42409

Qy	3060	GATGCTGACATGCTATTTTGTAGGGAGAAAAATATGTGCTAATGATATTTTGTAGTTAAA	3119
Db	42408	GATGCTGACATGCTATTTTGTAGGGAGAAAAATATGTGCTAATGATATTTTGTAGTTAAA	42349
Qy	3120	ATATCTTTTGGGAGGATTTGCTGAAAAAGTTGCACCTTTTGTACAAATGCTTATGCTTGGT	3179
Db	42348	ATATCTTTTGGGAGGATTTGCTGAAAAAGTTGCACCTTTTGTACAAATGCTTATGCTTGGT	42289
Qy	3180	ACAAGCTTATGCTGCTTAAATTTTAAAAAAATTAATAACTGTCTGTGAGAAACCAG	3239
Db	42288	ACAAGCTTATGCTGCTTAAATTTTAAAAAAATTAATAACTGTCTGTGAGAAACCAG	42230
Qy	3240	CTGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAAAATTACCTTTATATTCTAGTAT	3299
Db	42229	CTGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAAAATTACCTTTATATTCTAGTAT	42170
Qy	3300	TTTCAGCACTCCATAAATTTCTATTACCTAAATATGTCACACACTATTTTGTGATTTAAAAA	3359
Db	42169	TTTCAGCACTCCATAAATTTCTATTACCTAAATATGTCACACACTATTTTGTGATTTAAAAA	42110
Qy	3360	TTCTTACTAAGGAATAAAAACTTTTAATATACGATATGATATTGTCTAATAAATTAAAAAAG	3419
Db	42109	TTCTTACTAAGGAATAAAAACTTTTAATATACGATATGATATTGTCTAATAAATTAAAAAAG	42050
Qy	3420	ACATAATGGATGCTCAATTAGTTTTTAAGATATCTATACTATAGGGATACAAATCACTAC	3479
Db	42049	ACATAATGGATGCTCAATTAGTTTTTAAGATATCTATACTATAGGGATACAAATCACTAC	41990
Qy	3480	AGTTCTCAGATTTTACACCTTTTTTTTGTGTCATTTGCTGATGTCACACATTTTCCAATCTCT	3539
Db	41989	AGTTCTCAGATTTTACACCTTTTTTTTGTGTCATTTGCTGATGTCACACATTTTCCAATCTCT	41930
Qy	3540	TGCAAGCCTCCAGGCTCTGGCTTTGTCTACCTGCTCGTTCCCAATGATCTTAATGAAAA	3599
Db	41929	TGCAAGCCTCCAGGCTCTGGCTTTGTCTACCTGCTCGTTCCCAATGATCTTAATGAAAA	41870
Qy	3600	GTGCAAAAAGAAAAACCTACCAATTAA	3625
Db	41869	GTGCAAAAAGAAAAACCTACCAATTAA	41844

RESULT 5
BD176928
LOCUS BD176928 3274 bp DNA linear PAT 16-APR-2003
DEFINITION Novel bHLH type transcriptional gene, DEC2.
ACCESSION BD176928
VERSION BD176928.1 GI:30014187
KEYWORDS JP 2002300878-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3274)
AUTHORS Fujimoto,K., Shin,M. and Kato,Y.
TITLE Novel bHLH type transcriptional gene, DEC2
JOURNAL Patent: JP 2002300878-A 1 15-OCT-2002;
CHUGAI PHARMACEUT CO LTD
COMMENT OS Homo sapiens (human)
PN JP 2002300878-A/1
PD 15-OCT-2002
PF 19-AUG-1999 JP 1999233286
PI KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,G01N33/15,G01N33/50,G01N33/566//A61K31/711,A61K38/00, PC
A61K45/00,
PC A61K48/00,A61P43/00,C12P21/08,C12N15/00,C12N5/00,A61K37/02 CC
FH Key location/Qualifiers
FT CDS (135)..(1271).
FEATURES Location/Qualifiers
source 1..3274

/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"									
ORIGIN									
Query Match 62.9%; Score 2290; DB 6; Length 3274; Best local similarity 100.0%; Pred. No. 0; Matches 2290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1352	CGCCGCGCGCGCGCGTTC	CCCTGCTCCTCGGTGTTG	TGTCGCCCCCTCCCGAGAGCG	1411				
Db	985	CGCCGCGCGCGCGCGTTC	CCCTGCTCCTCGGTGTTG	TGTCGCCCCCTCCCGAGAGCG	1044				
QY	1412	GGGCGCGCGCGCGACCC	TCCTGCGCGACGAGTG	GGCGCCCTTGGGGCGCGCACCC	1471				
Db	1045	GGGCGCGCGCGCGACCC	TCCTGCGCGACGAGTG	GGCGCCCTTGGGGCGCGCACCC	1104				
QY	1472	CCAGCACCGCGACGGCC	GCACCCCTGCGCCCTT	CGCGCGCGCGAGCGGGGACCC	1531				
Db	1105	CCAGCACCGCGACGGCC	GCACCCCTGCGCCCTT	CGCGCGCGCGAGCGGGGACCC	1164				
QY	1532	GGAGAGCTCTGCTCAG	GAAGATCCCTCGCAGCC	AGGAAAGAGTCCCTGAATCCTTGC	1591				
Db	1165	GGAGAGCTCTGCTCAG	GAAGATCCCTCGCAGCC	AGGAAAGAGTCCCTGAATCCTTGC	1224				
QY	1592	GTCCCGAAGACGGAGG	TTCAAGCAGGTGAGAA	GTAAATACCTTTAAGGAGGTTCAA	1651				
Db	1225	GTCCCGAAGACGGAGG	TTCAAGCAGGTGAGAA	GTAAATACCTTTAAGGAGGTTCAA	1284				
QY	1652	GCAGAGTGAGAAGTT	AAAAATACCTTTAAG	GTCCTTTAAGGGAGGAAGTGTAATAGATGCAC	1711				
Db	1285	GCAGAGTGAGAAGTT	AAAAATACCTTTAAG	GTCCTTTAAGGGAGGAAGTGTAATAGATGCAC	1344				
QY	1712	GACAGGCATAAACA	GAACAACAACACAGG	TGTTATGTGTACATTCGGAGTTCCTGTTTT	1771				
Db	1345	GACAGGCATAAACA	GAACAACAACACAGG	TGTTATGTGTACATTCGGAGTTCCTGTTTT	1404				
QY	1772	GCTCATCCCGCACCC	CCCCCCTCCACACAC	TAACTCCCTTCTTCCCCCACCAGCT	1831				
Db	1405	GCTCATCCCGCACCC	CCCCCCTCCACACAC	TAACTCCCTTCTTCCCCCACCAGCT	1464				
QY	1832	GTAAAAGATCCTAT	CGGAAAGACACTGG	CTCTTTTAAATCCCCCAATAAATTTGC	1891				
Db	1465	GTAAAAGATCCTAT	CGGAAAGACACTGG	CTCTTTTAAATCCCCCAATAAATTTGC	1524				
QY	1892	CCCCTTTAGGCCAT	GTTCCTTAAATTTG	GAACCTAATTCGAGAGGAAGTAA	1951				
Db	1525	CCCCTTTAGGCCAT	GTTCCTTAAATTTG	GAACCTAATTCGAGAGGAAGTAA	1584				
QY	1952	GAAGGTCGTCTGT	GGCTGACATGGAA	CAGCAGGTAGTTGTTATGTAGAGCTAGTTCTC	2011				
Db	1585	GAAGGTCGTCTGT	GGCTGACATGGAA	CAGCAGGTAGTTGTTATGTAGAGCTAGTTCTC	1644				
QY	2012	CTTTGACGCTCTT	TGGAGTTGACAT	GGAAACAGCAGGTAGTTGTTATGTAGAGCTAGTTCTC	2071				
Db	1645	CTTTGACGCTCTT	TGGAGTTGACAT	GGAAACAGCAGGTAGTTGTTATGTAGAGCTAGTTCTC	1704				
QY	2072	AAAGCTGCCCTGC	CTGTTTAGGAGCG	GTTCACAAACAGATTGAGGCTCTTTTAAAGAT	2131				
Db	1705	AAAGCTGCCCTGC	CTGTTTAGGAGCG	GTTCACAAACAGATTGAGGCTCTTTTAAAGAT	1764				
QY	2132	TGAATTTACTCTT	CAATTTTCAATG	TTTCAAGGCGCATATATTTTCAAA	2191				
Db	1765	TGAATTTACTCTT	CAATTTTCAATG	TTTCAAGGCGCATATATTTTCAAA	1824				
QY	2192	GAAGTGAGGATGC	AGTTTCTCACGT	TGCAACCTATTCTGAAGTGGTTTAAATGGATCTC	2251				
Db	1825	GAAGTGAGGATGC	AGTTTCTCACGT	TGCAACCTATTCTGAAGTGGTTTAAATGGATCTC	1884				
QY	2252	TTAGTAACCTGCA	CTCGTTAAAGAA	ACACGGAGCTGGGCCATCGTCAGAACTAAGTCAGG	2311				
Db	1885	TTAGTAACCTGCA	CTCGTTAAAGAA	ACACGGAGCTGGGCCATCGTCAGAACTAAGTCAGG	1944				

QY	2312	GAAGGAGATGGATG	AGAAGGCCAGAA	TCAATTCCTAGTACATTTGCTAACACTTTATTGAG	2371
Db	1945	GAAGGAGATGGATG	AGAAGGCCAGAA	TCAATTCCTAGTACATTTGCTAACACTTTATTGAG	2004
QY	2372	AAATTGACCATGA	ATTAATGGACTCA	TCTTAATTTCTTCTAAGTCCATATATAGATAGAT	2431
Db	2005	AAATTGACCATGA	ATTAATGGACTCA	TCTTAATTTCTTCTAAGTCCATATATAGATAGAT	2064
QY	2432	ATCTATCTGTAC	AGATTTCTATTT	ATCCATAGATAGGTATCTATACATACACATCTCAAG	2491
Db	2065	ATCTATCTGTAC	AGATTTCTATTT	ATCCATAGATAGGTATCTATACATACACATCTCAAG	2124
QY	2492	TGCATCTATTC	CACTCTCATTA	TATCCATCATGTTCTTAAATTTTGTAAATCTTACTGTA	2551
Db	2125	TGCATCTATTC	CACTCTCATTA	TATCCATCATGTTCTTAAATTTTGTAAATCTTACTGTA	2184
QY	2552	AAAAAAGTGCA	CTGAACCTTCA	AAAAACAAAAACAACAACAACAACAACAAGTCC	2611
Db	2185	AAAAAAGTGCA	CTGAACCTTCA	AAAAACAAAAACAACAACAACAACAACAAGTCC	2244
QY	2612	AAACTGATATA	TATCCTATATT	CTGTGTTAAATTCAAAAGTGAACGAAAGCATTTAACTGGCC	2671
Db	2245	AAACTGATATA	TATCCTATATT	CTGTGTTAAATTCAAAAGTGAACGAAAGCATTTAACTGGCC	2304
QY	2672	AGTTTGTATT	GCAAAATGCT	GTAAGATATAGAACTCCTGTGAGGCTTCCTATCTC	2731
Db	2305	AGTTTGTATT	GCAAAATGCT	GTAAGATATAGAACTCCTGTGAGGCTTCCTATCTC	2364
QY	2732	CAAGTCTATG	TATTTCTGG	GAGACCAACCCAGATACCCAGATATACCAAGAAAGCTTTT	2791
Db	2365	CAAGTCTATG	TATTTCTGG	GAGACCAACCCAGATACCCAGATATACCAAGAAAGCTTTT	2424
QY	2792	TTAATAAGG	CTTAAACCA	AGACCTTGTCTAGATATTTTGTAGTTGTTGCCAAGGTAGCAC	2851
Db	2425	TTAATAAGG	CTTAAACCA	AGACCTTGTCTAGATATTTTGTAGTTGTTGCCAAGGTAGCAC	2484
QY	2852	TGTGAGAAAT	CTCACTTGG	ATGTAAGGGGTGAGACACACAGCTGACTATGAGT	2911
Db	2485	TGTGAGAAAT	CTCACTTGG	ATGTAAGGGGTGAGACACACAGCTGACTATGAGT	2544
QY	2912	GAGGAAAATA	TCTGGGTCT	TTTTCGTGAGTTTGGTGCATTTGCTGCTGCTTGTACTGT	2971
Db	2545	GAGGAAAATA	TCTGGGTCT	TTTTCGTGAGTTTGGTGCATTTGCTGCTGCTTGTACTGT	2604
QY	2972	TTGCCCTCAA	ACGCTGTG	TTTAAACACGTTAAACTCTTAGCCTACAAAGTGGCTCTTATG	3031
Db	2605	TTGCCCTCAA	ACGCTGTG	TTTAAACACGTTAAACTCTTAGCCTACAAAGTGGCTCTTATG	2664
QY	3032	TACATAGTT	GTTAATACAT	CCAATTAATGATGCTGACATGCTATTTTGTAGGGAGAA	3091
Db	2665	TACATAGTT	GTTAATACAT	CCAATTAATGATGCTGACATGCTATTTTGTAGGGAGAA	2724
QY	3092	ATATGTC	TAAATGAT	TTTTTGAGTTTAAATATCTTTTGGGAGGATTTGCTGAAAAGTTG	3151
Db	2725	ATATGTC	TAAATGAT	TTTTTGAGTTTAAATATCTTTTGGGAGGATTTGCTGAAAAGTTG	2784
QY	3152	CACTTTTG	TACAAATG	CTTATGCTTCAAGCTTATGCTGCTTAAATTTTAAAA	3211
Db	2785	CACTTTTG	TACAAATG	CTTATGCTTCAAGCTTATGCTGCTTAAATTTTAAAA	2844
QY	3212	AAATTAATA	TACTGTCT	GTGAGAAACAGCTGGTTTAGAAAAGTTTAGTATGTGACGATAA	3271
Db	2845	AAATTAATA	TACTGTCT	GTGAGAAACAGCTGGTTTAGAAAAGTTTAGTATGTGACGATAA	2904
QY	3272	ACTAGAA	ATTACCTTT	TATTTCTAGTATTTTTCAGCACTCCATAAATTTCTATTACCTAAAT	3331
Db	2905	ACTAGAA	ATTACCTTT	TATTTCTAGTATTTTTCAGCACTCCATAAATTTCTATTACCTAAAT	2964
QY	3332	ATTGCCACA	CTATTTTGT	GATTTTAAAAATTTCTTACTAAGGAATAAAACTTTTAATATACG	3391
Db	2965	ATTGCCACA	CTATTTTGT	GATTTTAAAAATTTCTTACTAAGGAATAAAACTTTTAATATACG	3024
QY	3392	ATATGATA	TATGTCTA	ATAATTAATAAAGACATAATGGATGCTCAATTAGTTTAAAGATAT	3451

Db 3025 ATATGATATTGTCTAATAATAAAGACATAAATGGATGCTCAATTAGTTTAAAGATAT 3084
Qy 3452 CTATAACTATAGGGATACAAATCACTACAGTTCTCAGATTTACACCTTTTGTGTCATT 3511
Db 3085 CTATAACTATAGGGATACAAATCACTACAGTTCTCAGATTTACACCTTTTGTGTCATT 3144
Qy 3512 GGCTTGATGTCACACATTTCCAAATCTCTTGCAAGCTCCAGGCTCTGGCTTTGTCTACCT 3571
Db 3145 GGCTTGATGTCACACATTTCCAAATCTCTTGCAAGCTCCAGGCTCTGGCTTTGTCTACCT 3204
Qy 3572 GCTCGTTCCCAATGTATCTTAATGAAAAGTGCAAAAGAAAACCTACCAATTAAGAAAAA 3631
Db 3205 GCTCGTTCCCAATGTATCTTAATGAAAAGTGCAAAAGAAAACCTACCAATTAAGAAAAA 3264
Qy 3632 AAAAAAAAAA 3641
Db 3265 AAAAAAAAAA 3274

RESULT 6
AX591648 AX591648 2304 bp DNA linear PAT 27-JAN-2003
DEFINITION Sequence 9 from Patent WO0246409.
ACCESSION AX591648
VERSION AX591648.1 GI:27950039

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Guo,X., Li,L., Patturajan,M., Shinkets,R.A., Casman,S.J.,
Malyankar,U.M., Tchernev,V.T., Vernet,C.A., Spytek,K.A.,
Shenoy,S.G., Alsobrook,J.P., Edinger,S., Peyman,J.A., Stone,D.J.,
Ellerman,K., Gangolli,E.A., Boldog,F.L., Colman,S.D., Eisen,A.J.,
Liu,X., Padigaru,M., Spaderna,S.K. and Zerhusen,B.D.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0246409-A 9 13-JUN-2002;
Curagen Corporation (US)

FEATURES
source
1. .2304
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 43.5%; Score 1583; DB 6; Length 2304;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACCTGAAGAGGGAGAGCGAGAGAGACTGGAGACGCACAGATCCCCCAAGGTCTC 60
Db 722 CTGCACCTGAAGAGGGAGAGCGAGAGAGACTGGAGACGCACAGATCCCCCAAGGTCTC 781
Qy 61 CCAAGCCTACCGTCCCACAGATTATTGTACAGAGCCCCCAAAATCGAAACAGAGGAAACG 120
Db 782 CCAAGCCTACCGTCCCACAGATTATTGTACAGAGCCCCCAAAATCGAAACAGAGGAAACG 841
Qy 121 AACAGCAGTTGAACATGGACGAAGGAATTCTCATTTGATATGTGTAAACCCAAAAGGAGCA 180
Db 842 AACAGCAGTTGAACATGGACGAAGGAATTCTCATTTGATATGTGTAAACCCAAAAGGAGCA 901
Qy 181 ATAGAGATTTTATAGGACTGGACTATTCTCTTTGTATATGTGTAAACCCAAAAGGAGCA 240
Db 902 ATAGAGATTTTATAGGACTGGACTATTCTCTTTGTATATGTGTAAACCCAAAAGGAGCA 961
Qy 241 TGAACGAGACGACACCAAGGATACCTACAAATTACCGCACAGATTATAGAAAAGAAA 300
Db 962 TGAACGAGACGACACCAAGGATACCTACAAATTACCGCACAGATTATAGAAAAGAAA 1021
Qy 301 GAAGAGACCGAATTAATGAATGCATTGCTCAGCTGAAAGATTTACTGCCTGAACATCTGA 360

Db 1022 GAAGAGACCGAATTAATGAATGCATTGCTCAGTGAAGATTTACTGCTGAACATCTGA 1081
Qy 361 AATTGACAACTCTGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAAACACT 420
Db 1082 AATTGACAACTCTGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAAACACT 1141
Qy 421 TAAAGAGCTTTAAACCGCTTAAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATG 480
Db 1142 TAAAGAGCTTTAAACCGCTTAAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATG 1201
Qy 481 GGGAGCGATCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGATTTC 540
Db 1202 GGGAGCGATCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGATTTC 1261
Qy 541 AAACATCGCCCAAGAAAGTCTTTGCAATACCTCTCCCGGTTTGAGAGCTGCACACCCAGGG 600
Db 1262 AAACATCGCCCAAGAAAGTCTTTGCAATACCTCTCCCGGTTTGAGAGCTGCACACCCAGGG 1321
Qy 601 AGCCGCGGTGTGTCCAGCTGATCAACCACTTGTGACGCCGTGGCCACCCAGTTCTTGCCCA 660
Db 1322 AGCCGCGGTGTGTCCAGCTGATCAACCACTTGTGACGCCGTGGCCACCCAGTTCTTGCCCA 1381
Qy 661 CCCGCGAGCTGTTGACTCAACAGGTCCCTCTGAGCAAAAGCACCGGCGTCCCTCGGCCG 720
Db 1382 CCCGCGAGCTGTTGACTCAACAGGTCCCTCTGAGCAAAAGCACCGGCGTCCCTCGGCCG 1441
Qy 721 CCGGTCGCGCGGCCCTGCTGAGCGCGCGGGGCGAGAAGCTGGAGCCCCCTCGCCT 780
Db 1442 CCGGTCGCGCGGCCCTGCTGAGCGCGCGGGGCGAGAAGCTGGAGCCCCCTCGCCT 1501
Qy 781 ACTGCGTGCCCGTATCCAGCGGACTCAGCCACGCGCGAGCTCGCCGCGGAGAACGACA 840
Db 1502 ACTGCGTGCCCGTATCCAGCGGACTCAGCCACGCGCGAGCTCGCCGCGGAGAACGACA 1561
Qy 841 CGGACACCGACAGCGGCTACGGCGGAAGCGAGGCCCGGCGGACCGCGAGAAAGGCA 900
Db 1562 CGGACACCGACAGCGGCTACGGCGGAAGCGAGGCCCGGCGGACCGCGAGAAAGGCA 1621
Qy 901 AAGCGCGGGGCGAGCGCGTCAACCATCAAGCAGGAGCCTCCCGGAGGACTCGCCGG 960
Db 1622 AAGCGCGGGGCGAGCGCGTCAACCATCAAGCAGGAGCCTCCCGGAGGACTCGCCGG 1681
Qy 961 CGCCCAAGAGGATGAAGCTGGATTCCCGCGGCGGCGGAGCGCGGCCCGGGGGCG 1020
Db 1682 CGCCCAAGAGGATGAAGCTGGATTCCCGCGGCGGCGGAGCGCGGCCCGGGGGCG 1741
Qy 1021 GCGCGGCGGCGGCGGCGGCGGCTTCTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Db 1742 GCGCGGCGGCGGCGGCGGCGGCTTCTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCG 1801
Qy 1081 TGAGACCCGACGCGCGCTGCTCAGCTCGCTGTGGCGGCTTGGCGGAGGCGGCGGCGG 1140
Db 1802 TGAGACCCGACGCGCGCTGCTCAGCTCGCTGTGGCGGCTTGGCGGAGGCGGCGGCGG 1861
Qy 1141 CCTTCCCGCAGCCCGCGCGCGGCGGCGGCTTCTGCTGCGGCTTCTGCTTCTCTCGC 1200
Db 1862 CCTTCCCGCAGCCCGCGCGCGGCGGCGGCTTCTGCTGCGGCTTCTGCTTCTCTCGC 1921
Qy 1201 CTTCTGCAGCTCCCGCTACGTGACGCCCTTCTTGGACAAGAGCGGCTTGGAGAAGTATC 1260
Db 1922 CTTCTGCAGCTCCCGCTACGTGACGCCCTTCTTGGACAAGAGCGGCTTGGAGAAGTATC 1981
Qy 1261 TGTACCCGCGGCGGCTGCGCGCGGCTTCCCGCTGTATACCCCGGCTATCCCGGCGG 1320
Db 1982 TGTACCCGCGGCGGCTGCGCGCGGCTTCCCGCTGTATACCCCGGCTATCCCGGCGG 2041
Qy 1321 CGGAGCGCGGCGGAGCGCGCGGCTGCGCGCGGCGGCGGCGGCTTCCCTGCTGT 1380
Db 2042 CGGAGCGCGGCGGAGCGCGCGGCTGCGCGCGGCGGCGGCGGCTTCCCTGCTGT 2101
Qy 1381 CCTCGGTGTTGTGCGCCCTCCCGAGAAAGCGGGCGGCGGCGGCGGCGGCTTCCCTGCTGT 1440
Db 2102 CCTCGGTGTTGTGCGCCCTCCCGAGAAAGCGGGCGGCGGCGGCGGCGGCTTCCCTGCTGT 2161

```
QY 1441 ACGAGTGGCGCCCTTGGGGCGCGCACCCCGCAGCACCCCGCACCGCCGCGCACCCACCTGC 1500
|||||
Db 2162 ACGAGTGGCGCCCTTGGGGCGCGCACCCCGCAGCACCCCGCACCGCCGCGCACCCACCTGC 2221
|||||
QY 1501 CCTTCGCGCGCGCGCGAGCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 1560
|||||
Db 2222 CCTTCGCGCGCGCGCGAGCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 2281
|||||
QY 1561 AGCCAGGAAAGGAAAGCTCCCTGA 1583
|||||
Db 2282 AGCCAGGAAAGGAAAGCTCCCTGA 2304
|||||

RESULT 7
BD011867
LOCUS BD011867 1511 bp DNA linear PAT 02-AUG-2002
DEFINITION Novel bHLH type transcriptional gene, DEC2.
ACCESSION BD011867
VERSION BD011867.1 GI:22092056
KEYWORDS WO 0114551-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1511)
Fujimoto,K., Shin,M. and Kato,Y.
Novel bHLH type transcriptional gene, DEC2
Patent: WO 0114551-A 10 01-MAR-2001;
CHUGAI PHARM CO LTD,KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
COMMENT OS Homo sapiens (human)
PN WO 0114551-A/10
PD 01-MAR-2001
PF 19-JUN-2000 WO 2000JP003991
PR 19-AUG-1999 JP 99F 233286
PI KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
PC C12N15/12,C07L14/475,C07K16/18
CC

FEATURES
source
1. .1511
Location/Qualifiers
(2) . .(1453).
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 37.8%; Score 1378; DB 6; Length 1511;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 GATACCTACAAATTACCGCACAGATTAAAGAAAGAAAGAGACCGAATTAATGAA 320
|||||
Db 134 GATACCTACAAATTACCGCACAGATTAAAGAAAGAAAGAGACCGAATTAATGAA 193
|||||
QY 321 TGCATTGCTCAGCTGAAAGATTTTACTGCTGAACATCTGAAATTGACAACTCTGGGACAT 380
|||||
Db 194 TGCATTGCTCAGCTGAAAGATTTTACTGCTGAACATCTGAAATTGACAACTCTGGGACAT 253
|||||
QY 381 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTAACCGCCTTA 440
|||||
Db 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTAACCGCCTTA 313
|||||
QY 441 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATGGGAGCGCATCTCTGAAATCG 500
|||||
Db 314 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAAATGGGAGCGCATCTCTGAAATCG 373
|||||
QY 501 CCCATTACGCCGACTTGGATGCGTTCCACTCGGGAATTTCAAACATGCGCCAAAGAGTC 560
|||||
Db 374 CCCATTACGCCGACTTGGATGCGTTCCACTCGGGAATTTCAAACATGCGCCAAAGAGTC 433
|||||
QY 561 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTGTCCAGCTG 620
|||||
```

RESULT 8
AK094422

```
Db 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCCGGTTGTGTCCAGCTG 493
QY 621 ATCAACCACTTGCACGCGCTGGCCACCCAGTTCTTTGCCACCCCGCAGCTGTGACTCAA 680
|||||
Db 494 ATCAACCACTTGCACGCGCTGGCCACCCAGTTCTTTGCCACCCCGCAGCTGTGACTCAA 553
|||||
QY 681 CAGGTCCCTCTGAGCAAAAGGACCCGGCGCTCCCTCGGCCCGCGGGTCCGGCGCGCGCCCC 740
|||||
Db 554 CAGGTCCCTCTGAGCAAAAGGACCCGGCGCTCCCTCGGCCCGCGGGTCCGGCGCGCGCCCC 613
|||||
QY 741 TGCCTGGAGCGCGCGGGGAGAAAGCTGGAGCCCCCTCGCCTACTGCTGCCCCGTATCCAG 800
|||||
Db 614 TGCCTGGAGCGCGCGGGGAGAAAGCTGGAGCCCCCTCGCCTACTGCTGCCCCGTATCCAG 673
|||||
QY 801 CGGACTCAGCCCCAGCGCCGAGCTCGCCGCCGAGAAACGACACGGACACCGACAGCGGTAC 860
|||||
Db 674 CGGACTCAGCCCCAGCGCCGAGCTCGCCGCCGAGAAACGACACGGACACCGACAGCGGTAC 733
|||||
QY 861 GCGGGGAAAGCCGAGGCCCCGGCCGACCCGAGAGAAAGGCAAAAGGCGGGGGCGAGCCGC 920
|||||
Db 734 GCGGGGAAAGCCGAGGCCCCGGCCGACCCGAGAGAAAGGCAAAAGGCGGGGGCGAGCCGC 793
|||||
QY 921 GTACCATCAAGCAGGAGCCTCCCGGGAGAGCTCGCCGGCGCCCAAGAGGATGAAGCTG 980
|||||
Db 794 GTACCATCAAGCAGGAGCCTCCCGGGAGAGCTCGCCGGCGCCCAAGAGGATGAAGCTG 853
|||||
QY 981 GATTCCCGCGCGCGGCGAGCGCGCGCCCGCGGGGGCGCGCGCGCGCGCGCGCGAGCC 1040
|||||
Db 854 GATTCCCGCGCGCGGCGAGCGCGCGCGCGCGGGGGCGCGCGCGCGCGCGCGCGAGCC 913
|||||
QY 1041 GCGCTTCTGGGGCCCGACCCCTGCGCCCGCGCGCGCGCTGCTGAGACCCCGCGCGCCCTG 1100
|||||
Db 914 GCGCTTCTGGGGCCCGACCCCTGCGCCCGCGCGCGCGCTGCTGAGACCCCGCGCGCCCTG 973
|||||
QY 1101 CTCAGCTCGCTGTGGCGTTTCGGCGGAGGCGGAGCGCGCCCTTCCCGCAGCCCGCGGCC 1160
|||||
Db 974 CTCAGCTCGCTGTGGCGTTTCGGCGGAGGCGGAGCGCGCCCTTCCCGCAGCCCGCGGCC 1033
|||||
QY 1161 GCGCGGCGCCCTTCTGCTGCCCTTCTGCTTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCT 1220
|||||
Db 1034 GCGCGGCGCCCTTCTGCTGCCCTTCTGCTTCTCTGCTTCTCTGCTGCTGCTGCTGCTGCT 1093
|||||
QY 1221 GTGACGCCCTTCTGGAACAAGAGCGCGCTGGAGAAAGTATCTGTACCCGCGCGCGGTGCC 1280
|||||
Db 1094 GTGACGCCCTTCTGGAACAAGAGCGCGCTGGAGAAAGTATCTGTACCCGCGCGCGGTGCC 1153
|||||
QY 1281 GCGCGGTTCCGCTGTATACCCCGGATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 1340
|||||
Db 1154 GCGCGGTTCCGCTGTATACCCCGGATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 1213
|||||
QY 1341 GCGCGGTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 1400
|||||
Db 1214 GCGCGGTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 1273
|||||
QY 1401 CCGGAGAAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 1460
|||||
Db 1274 CCGGAGAAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 1333
|||||
QY 1461 GCGCGGCAACCCCGCAGCACCCGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1520
|||||
Db 1334 GCGCGGCAACCCCGCAGCACCCGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1393
|||||
QY 1521 CCGGGGAAACCCCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCCGAGAAAGAGCTCCC 1580
|||||
Db 1394 CCGGGGAAACCCCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCCGAGAAAGAGCTCCC 1453
|||||
QY 1581 TGAATCCTTGGTCCCGAAGGACCGAGGTTCAAGCAGAGTGAGAGTTAAATACCT 1638
|||||
Db 1454 TGAATCCTTGGTCCCGAAGGACCGAGGTTCAAGCAGAGTGAGAGTTAAATACCT 1511
|||||
```


LOCUS AK094422 1973 bp mRNA linear PRI 30-JAN-2004
DEFINITION Homo sapiens cDNA FLJ37103 fis, clone BRACE2019327.
ACCESSION AK094422
VERSION AK094422.1 GI:21753480
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS

Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 1973)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE

NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1. .1973
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source

/db_xref="taxon:9606"
/clone="BRACE2019327"
/tissue_type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: pME18SFL3"

ORIGIN
Query Match 25.3%; Score 921; DB 9; Length 1973;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2533 TTTTGTAACTCTTACTGTAAAAAAGTGCACCTTCAAAACAAAAACAACAAAC 2592
Db 1 TTTTGTAACTCTTACTGTAAAAAAGTGCACCTTCAAAACAAAAACAACAAAC 60

QY 2593 AACACAAAAAACAAGTCAAACTGATATATCCTATATCTGTTAAATTCAAAAGTGAA 2652
Db 61 AACACAAAAAACAAGTCAAACTGATATATCCTATATCTGTTAAATTCAAAAGTGAA 120

QY 2653 CGAAAGCATTTAACTGGCCAGTTTGTATTGCAATGCTGTAAAGATATAGAATGAAGTCC 2712
Db 121 CGAAAGCATTTAACTGGCCAGTTTGTATTGCAATGCTGTAAAGATATAGAATGAAGTCC 180

QY 2713 TGTGAGGCCTTCCTATCTCAAGTCTATGTATTTTCTGGAGACCAACCCAGATACCAGAT 2772
Db 181 TGTGAGGCCTTCCTATCTCAAGTCTATGTATTTTCTGGAGACCAACCCAGATACCAGAT 240

QY 2773 AATCACAAGAAAGCTTTTAAAGGCTTAAACCAAGACCTTGTCTAGATATTTTAG 2832
Db 241 AATCACAAGAAAGCTTTTAAAGGCTTAAACCAAGACCTTGTCTAGATATTTTAG 300

QY 2833 TTTGTTGCCAAGGTAGCACTGTGAGAAATCTCCTCTGGATGTTATGTAAGGGGTGAGACA 2892
Db 301 TTTGTTGCCAAGGTAGCACTGTGAGAAATCTCCTCTGGATGTTATGTAAGGGGTGAGACA 360

QY 2893 CAACAGTCTGACTATGAGTGAGGAAAAATATCTGGTCTTTTCGTCAGTTTGGTGCATTG 2952
Db 361 CAACAGTCTGACTATGAGTGAGGAAAAATATCTGGTCTTTTCGTCAGTTTGGTGCATTG 420

QY 2953 CTGCTGCTGTTGCTACTGTTTGCCTCAAAACGCTGTGTTAAACAACCGTTAAACTCTTAGC 3012
Db 421 CTGCTGCTGTTGCTACTGTTTGCCTCAAAACGCTGTGTTAAACAACCGTTAAACTCTTAGC 480

QY 3013 CTACAAGGTGGCTCTTATGTACATAGTTGTTTAATACATCCAAATTAATGATGCTGACATG 3072
Db 481 CTACAAGGTGGCTCTTATGTACATAGTTGTTTAATACATCCAAATTAATGATGCTGACATG 540

QY 3073 CTATTTTGTAGGGAGAAATATGTGCTAATGATATTTTGTAGTTAAATAATCTTTTGGGG 3132
Db 541 CTATTTTGTAGGGAGAAATATGTGCTAATGATATTTTGTAGTTAAATAATCTTTTGGGG 600

QY 3133 AGGATTTGCTGAAAAAGTTGCACTTTTGTACAAATGCTTATGCTTGGTACAAGTTATGCT 3192
Db 601 AGGATTTGCTGAAAAAGTTGCACTTTTGTACAAATGCTTATGCTTGGTACAAGTTATGCT 660

QY 3193 GTCTTAAATTTATTTTAAAAAATAAATACTGTCTGTGAGAAACCCAGCTGGTTTAAAAA 3252
Db 661 GTCTTAAATTTATTTTAAAAAATAAATACTGTCTGTGAGAAACCCAGCTGGTTTAAAAA 719

QY 3253 GTTTAGTATGTGACGATAAACTAGAAATTAACCTTTATATTCTAGTATTTTCAGCACTCCA 3312
Db 720 GTTTAGTATGTGACGATAAACTAGAAATTAACCTTTATATTCTAGTATTTTCAGCACTCCA 779

QY 3313 TAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTTAAAAATTTCTTAAAGGA 3372
Db 780 TAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTTTAAAAATTTCTTAAAGGA 839

QY 3373 ATAAAAACCTTTAATATACGATATGATATTGTCTAATAATTAATAAGACATAATGGATGC 3432
Db 840 ATAAAAACCTTTAATATACGATATGATATTGTCTAATAATTAATAAGACATAATGGATGC 899

QY 3433 TCAATTAGTTTTTAAGATATCTATAACTATAGGGATACAAATCACTACAGTTCTCAGATTT 3492

Db 900 TCAATTAGTTTTAAGATATCTATAACTATAGGATACAAAATCACTACAGTTCTCAGATT 959
QY 3493 ACACCTTTTTTTTGTCTATTGGCTTGATGTCACACATTTCCAAATCTCTTGCAAGCCTCCAG 3552
Db 960 ACAGCTTTTTTTTGTCTATTGGCTTGATGTCACACATTTCCAAATCTCTTGCAAGCCTCCAG 1019
QY 3553 GCTCTGGCTTTGTCTACCTGCTCGTTCCCAATGTATCTTAATGAAAAAGTGCAAAAGAAA 3612
Db 1020 GCTCTGGCTTTGTCTACCTGCTCGTTCCCAATGTATCTTAATGAAAAAGTGCAAAAGAAA 1079
QY 3613 ACCTACCAATTAA 3625
Db 1080 ACCTACCAATTAA 1092

RESULT 9
CQ434066
LOCUS CQ434066 753 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 19100 from Patent WO0151628.
ACCESSION CQ434066
VERSION CQ434066.1 GI:41386295
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE Novel genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 19100 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
1. 753
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 17.3%; Score 629; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2445 GATTTCTATTTATCCATAGATAGGTATCTATACATACACATCTCAAGTGCACTATTTCCC 2504
Db 62 GATTTCTATTTATCCATAGATAGGTATCTATACATACACATCTCAAGTGCACTATTTCCC 121
QY 2505 ACTCTCAATTAATCCATCATGTTCCCTAAATTTTGTAACTTTTGAATCTTACTGTAAAAAAGTGCAC 2564
Db 122 ACTCTCAATTAATCCATCATGTTCCCTAAATTTTGTAACTTTTGAATCTTACTGTAAAAAAGTGCAC 181
QY 2565 TGAACCTCAAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 2624
Db 182 TGAACCTCAAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 241
QY 2625 CTATATCTGTATAAATTCAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTGTATTGCA 2684
Db 242 CTATATCTGTATAAATTCAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTGTATTGCA 301
QY 2685 AATGCTGTAAAGATATAGAAATGAAGTCCTGTGAGGCCCTTCCCTATCTCCAAGTCTATGTAT 2744
Db 302 AATGCTGTAAAGATATAGAAATGAAGTCCTGTGAGGCCCTTCCCTATCTCCAAGTCTATGTAT 361
QY 2745 TTTCTGGAGACCAAAACCAAGATACAGATAATCACAAGAAAGCTTTTTTAATAAGGCTTA 2804
Db 362 TTTCTGGAGACCAAAACCAAGATACAGATAATCACAAGAAAGCTTTTTTAATAAGGCTTA 421
QY 2805 AACCAAGACCTTGTCTAGATATTTTAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTC 2864
Db 422 AACCAAGACCTTGTCTAGATATTTTAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTC 481
QY 2865 ACTTGGATGTTATGTAAGGGTGAGACACAACAGTCTGACTATGAGTGAGGAAATATCT 2924

Db 482 ACTTGGATGTTATGTAAGGGGTGAGACACACACAGTCTGACTATGAGTGAGGAAAATATCT 541
QY 2925 GGGTCTTTTTTCGTCACTTTGGTGCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2984
Db 542 GGGTCTTTTTTCGTCACTTTGGTGCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
QY 2985 TGTGTTTAAACAACGTTAAACTCTTAGCCCTACAAAGGTGGCTCTTATGTACATAGTTGTTA 3044
Db 602 TGTGTTTAAACAACGTTAAACTCTTAGCCCTACAAAGGTGGCTCTTATGTACATAGTTGTTA 661
QY 3045 ATACATCCAAATTAATGATGTCTGACATGC 3073
Db 662 ATACATCCAAATTAATGATGTCTGACATGC 690

RESULT 10
CQ727066
LOCUS CQ727066 909 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 13000 from Patent WO02068579.
ACCESSION CQ727066
VERSION CQ727066.1 GI:42291946
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 13000 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
1. 909
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 15.2%; Score 552; DB 6; Length 909;
Best Local Similarity 100.0%; Pred. No. 7.7e-287;
Matches 552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 369 ACTCTGGGACATCTGGAGAAAAGCTGTAGTCTTGGAAATTAACCTTGAAACACCTTAAAGCT 428
Db 181 ACTCTGGGACATCTGGAGAAAAGCTGTAGTCTTGGAAATTAACCTTGAAACACCTTAAAGCT 240
QY 429 TTAACCGCCTTAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAAATGGGAGCGA 488
Db 241 TTAACCGCCTTAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAAATGGGAGCGA 300
QY 489 TCTCTGAAATCGCCCATTCAGTCCGACTTGGATCGCTTCCACTCGGGATTTCAAACATGC 548
Db 301 TCTCTGAAATCGCCCATTCAGTCCGACTTGGATCGCTTCCACTCGGGATTTCAAACATGC 360
QY 549 GCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGG 608
Db 361 GCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGG 420
QY 609 TGTGTCCAGCTGATCAACCACCTTGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 668
Db 421 TGTGTCCAGCTGATCAACCACCTTGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 480
QY 669 CTGTTGACTCAACAGGTCCCTCTGAGCAAAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 728
Db 481 CTGTTGACTCAACAGGTCCCTCTGAGCAAAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 540
QY 729 GCGGCCGCCCTTGCCTGGAGCGCGCGGGGCGAGAAAGCTGGAGCCCCCTCGCCCTACTGCGTG 788
Db 541 GCGGCCGCCCTTGCCTGGAGCGCGCGGGGCGAGAAAGCTGGAGCCCCCTCGCCCTACTGCGTG 600
QY 789 CCGGTCTATCCAGCGGACTCAGCCCGCGCGGAGCTCGCCCGCGGAGACGACACGAGACACC 848

Db 601 CCGTCATCCAGCGGACTCAGCCAGCGCGAGCTCGCCGCGAGAACGACACGACACC 660

Qy 849 GACAGCGGCTACGGCGGCGAAGCCGAGGCGCGCGCGGACCGCGAGAAAGGCAAGGCGCG 908

Db 661 GACAGCGGCTACGGCGGCGAAGCCGAGGCGCGCGGACCGCGAGAAAGGCAAGGCGCG 720

Qy 909 GGGCGGAGCCGC 920

Db 721 GGGCGGAGCCGC 732

RESULT 11

HSA270695

LOCUS HSA270695 593 bp mRNA linear PRI 30-MAY-2001

DEFINITION Homo sapiens partial unknown mRNA from drug-resistant melanoma cells, 3'UTR, clone DMS-11.

ACCESSION AJ270695

VERSION AJ270695.1 GI:8017389

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Grottko,C., Mantwill,K., Dietel,M., Schadendorf,D. and Lage,H. Identification of differentially expressed genes in human melanoma cells with acquired resistance to various antineoplastic drugs

AUTHORS Int. J. Cancer 88 (4), 535-546 (2000)

TITLE

JOURNAL

MEDLINE 20514193

PUBMED 11058868

REFERENCE 2 (bases 1 to 593)

AUTHORS Lage,H.

TITLE Direct Submission

JOURNAL Submitted (18-OCT-1999) Lage H., Institute of Pathology, Humboldt University Berlin, University Hospital Charite, Schumannstr. 20/21, D-10117 Berlin, GERMANY

FEATURES

Location/Qualifiers

1..593

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DSM-11"

/cell_line="MeWo"

/cell_type="melanoma"

<1..593

3'UTR

polyA_site 578

ORIGIN

Query Match 12.6%; Score 460; DB 9; Length 593;

Best Local Similarity 99.8%; Pred. No. 5.7e-237;

Matches 580; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1978 GGTGAACCCCGGGGTAGGGGAAGATGTTAACACCTTTGACGCTCTTTGGAGTTGACATGG 2037

Db 7 GGTGAACCCCGGGGTAGGGGAAGATGTTAACACCTTTGACGCTCTTTGGAGTTGACATGG 66

Qy 2038 AACAGCAGGTAGTTGTTATGTAGCTAGTTCTCAAAGCTGCCCTGCCTGTTTAGGAGG 2097

Db 67 AACAGCAGGTAGTTGTTATGTAGCTAGTTCTCAAAGCTGCCCTGCCTGTTTAGGAGG 126

Qy 2098 CGTTCCCAAAACAGATTGAGGCTCTTTTAGAATTTGAATTTACTCTTTCAGTATTTCTAA 2157

Db 127 CGTTCCCAAAACAGATTGAGGCTCTTTTAGAATTTGAATTTACTCTTTCAGTATTTCTAA 185

Qy 2158 TGTTTCAGCTTTCTAAAGGCATATATTTTCAAAGAGTGAGGATGCAGTTTCTCACGTT 2217

Db 186 TGTTTCAGCTTTCTAAAGGCATATATTTTCAAAGAGTGAGGATGCAGTTTCTCACGTT 245

Qy 2218 GCAACCTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAATTTGCACCTCGTTAAAGAAA 2277

Db 246 GCAACCTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAATTTGCACCTCGTTAAAGAAA 305

Qy 2278 CACGAGCTGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAA 2337

Db 306 CACGAGCTGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAA 365

Qy 2338 TCATTCCCTAGTACATTTCCTAACACCTTTATTGAGAAATTGACCATGAATTAATGGACTCA 2397

Db 366 TCATTCCCTAGTACATTTCCTAACACCTTTATTGAGAAATTGACCATGAATTAATGGACTCA 425

Qy 2398 TCCTTAATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTAT 2457

Db 426 TCCTTAATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTAT 485

Qy 2458 CCATAGATAGGTATCTATATACATACACATCTCAAGTGCATCTATTTCCCACTCTCATTAATC 2517

Db 486 CCATAGATAGGTATCTATATACATACACATCTCAAGTGCATCTATTTCCCACTCTCATTAATC 545

Qy 2518 CATCATGTTCTCTAAATTTTGTAACTTCTTACTGTATATAAAAAA 2558

Db 546 CATCATGTTCTCTAAATTTTGTAACTTCTTACTGTATATAAAAAA 586

RESULT 12

BV208394

LOCUS BV208394 797 bp DNA linear STS 01-JUL-2004

DEFINITION BHLHB3 2085 Rhesus macaque genomic DNA Macaca mulatta STS genomic clone MMA2085, sequence tagged site.

ACCESSION BV208394

VERSION BV208394.1 GI:49533077

KEYWORDS STS.

SOURCE Macaca mulatta (rhesus monkey)

ORGANISM Macaca mulatta

REFERENCE 1 (bases 1 to 797)

AUTHORS Spindel,E.R., Pauley,M., Jia,Y., Boyle,N., Jiang,S., Gravett,C., Lupo,S.L., Ali,H., Ojeda,S.R. and Norrgren,R.B. Targeted amplification of the 3' end of rhesus macaque orthologs of human genes

TITLE human genes

JOURNAL Unpublished (2004)

COMMENT

Contact: Spindel ER

Division of Neuroscience

Oregon National Primate Research Center

505 NW 185th Avenue, Beaverton, OR 97006, USA

Tel: 403-690-5388

Fax: 503-690-5384

Email: spindele@ohsu.edu

Primer A: atgaattaatgactcatctt

Primer B: actaaacttttctaaaccagc

STS size: 797

PCR Profile:

Hot Start: 95 degrees C for 2.00 min

Denaturation: 95 degrees C for 0.50 min

Annealing: 51 degrees C for 0.50 min

Polymerization: 72 degrees C for 1.00 min

PCR Cycles: 35

Extension: 72 degrees C for 7.0 min

Thermal Cycler: MJ Instruments PTC100

Protocol:

Template: 200 ng

Primer: each 1uM

dNTP's: each 200 uM

Taq Polymerase: 0.05 units/ul (Fast Start High Fidelity, Roche)

Total Vol: 50 ul

Buffer:

MgCl2: 1.8 mM

Fast Start polymerase reaction buffer (Roche)

Bases 156-797 are 97% homologous (Blast) to bases 2596-3238 of NM_030762.1. Primers were chosen to amplify genomic DNA in the 3'

region of BHLHB3. As human sequence was used to design the primers, the primer sequences are not included in the rhesus sequence provided below. To obtain additional information regarding primers or clones contact: Dr. Robert Norgren; Dept of Genetics, Cell Biology & Anatomy; University of Nebraska Medical Center; 986395 Nebraska Medical Center; Omaha, NE 68198. Email: rnorgren@unmc.edu

A database containing sequences associated with this project can be found at: <http://rhesusgenechip.unomaha.edu/index.html>.

FEATURES
source
1. .797
/organism="Macaca mulatta"
/mol_type="genomic DNA"
/strain="Indian origin"
/db_xref="taxon:9544"
/clone="MMA2085"
/clone_lib="Rhesus macaque genomic DNA"
/dev_stage="Adult"
/note="Organ: Liver; Vector: pGEM-T Easy; V-type: Plasmid; cDNA amplified from rhesus genomic DNA with the human forward and reverse primers listed above and subcloned into pGEM-T Easy."
3
/gene="BHLHB3"
/note="basic helix-loop-helix domain containing, class B, 3"
<1. .>797
/gene="BHLHB3"

Query Match 11.3%; Score 413; DB 11; Length 797;
Best Local Similarity 99.8%; Pred. No. 1.8e-211;
Matches 463; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2702 GAATGAAGTCCTGTGAGGCCCTTCCTATCTCCAAAGTCTATGTATTTCTGGAGACCAAACC 2761
Db
QY 2762 AGATACCAGATAATCAAAAGAAAGCTTTTAAAGGCTTAAGGCTTAACCAAGACCTTGCTA 2821
Db 322 AGATACCAGATAATCAAAAGAAAGCTTTTAAAGGCTTAACCAAGACCTTGCTA 381
QY 2822 GATATTTTAGTTGTGTCGAAGGTAGCACTGTGAGAAATCTCACTGGATGTTATGTAA 2881
Db 382 GATATTTTAGTTGTGTCGAAGGTAGCACTGTGAGAAATCTCACTGGATGTTATGTAA 441
QY 2882 GGGGTGAGACACACAGTCTGACTATGAGTGAGGAAATATCTGGGTCTTTTCGTCAGTT 2941
Db 442 GGGGTGAGACACACAGTCTGACTATGAGTGAGGAAATATCTGGGTCTTTTCGTCAGTT 501
QY 2942 TGGTGCAATTTGCTGCTGTGTTGCTACTGTTTCCTCAAACGCTGTGTTAAACAACGTT 3001
Db 502 TGGTGCAATTTGCTGCTGTGTTGCTACTGTTTCCTCAAACGCTGTGTTAAACAACGTT 561
QY 3002 AAACCTCTTAGCTACAAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATGA 3061
Db 562 AAACCTCTTAGCTACAAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATGA 621
QY 3062 TGTCTGACATGCTATTTTGTAGGGAGAAATATGTGCTAATGATATTTTGAGTTAAAT 3121
Db 622 TGTCTGACATGCTATTTTGTAGGGAGAAATATGTGCTAATGATATTTTGAGTTAAAT 681
QY 3122 ATCTTTTGGGAGGATTTGCTGAAAGTTGCACCTTTTGTACAA 3165
Db 682 ATCTTTTGGGAGGATTTGCTGAAAGTTGCACCTTTTGTACAA 725

RESULT 13
CQ424141
LOCUS CQ424141 452 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 9175 from Patent WO0151628.
ACCESSION CQ424141
VERSION CQ424141.1 GI:41376370

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.
TITLE Novel genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 9175 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. .462
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 10.7%; Score 391; DB 6; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.4e-199;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2445 GATTTCATTATCCATAGATAGGTATCTATACATACACATCTCAAGTGCATCTATCCC 2504
Db 15 GATTTCATTATCCATAGATAGGTATCTATACATACACATCTCAAGTGCATCTATCCC 74
QY 2505 ACTCTCATTAATCCATCATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGCAC 2564
Db 75 ACTCTCATTAATCCATCATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGCAC 134
QY 2565 TGAACCTTCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2624
Db 135 TGAACCTTCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 194
QY 2625 CTATATTTCTGTTAAATTCAAAGTGAAACGAAAGCATTTAACTGGCCAGTTTGTATGCA 2684
Db 195 CTATATTTCTGTTAAATTCAAAGTGAAACGAAAGCATTTAACTGGCCAGTTTGTATGCA 254
QY 2685 AATGCTGTAAAGATATAGAATGAAGTCCTGTGAGGCCTTCCTATCTCCAAGTCTATGTAT 2744
Db 255 AATGCTGTAAAGATATAGAATGAAGTCCTGTGAGGCCTTCCTATCTCCAAGTCTATGTAT 314
QY 2745 TTTCTGGAGACCAACACAGATACCAGATATATCAACAAAGAAAGCTTTTAAATAAGGCTTA 2804
Db 315 TTTCTGGAGACCAACACAGATACCAGATATATCAACAAAGAAAGCTTTTAAATAAGGCTTA 374
QY 2805 AACCAAGACCTTGTCTAGATATTTTGTAGTTT 2835
Db 375 AACCAAGACCTTGTCTAGATATTTTGTAGTTT 405

RESULT 14
CQ432989
LOCUS CQ432989 447 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 18023 from Patent WO0151628.
ACCESSION CQ432989
VERSION CQ432989.1 GI:41385218

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.
TITLE Novel genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 18023 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. .447
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

1 TGGAGACCAAC CAGATAC CAGATAAT CACAAAGAAAGCTTTT TAATAAGGCTTAAACC 180

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 17, 2004, 04:12:41 : Search time 4563.79 Seconds
(without alignments)
12064.617 Million cell updates/sec

Title: US-10-078-650-11
Perfect score: 1511
Sequence: 1 catggacgaaggaattcctc.....gtgagaagttaataaccct 1511

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 32822875 seqs, 18219865908 residues

Word size : 0

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484	32.0	523	1	AI672308 ty63f03.x
2	434	28.7	715	4	BM542244 AGENCOURT
3	431	28.5	522	1	AI700204 wd41h01.x
4	430	28.5	557	1	AI819798 wj45a04.x
5	387	25.6	526	1	AI459114 tj65c02.x
6	380	25.1	515	2	BE857082 7g23f11.x
7	380	25.1	557	2	AW073334 wy97g05.x
8	322	21.3	434	1	AA996006 os26c04.s
9	310	20.5	310	9	CG464925 KRIBB 2D
10	301	19.9	301	1	AI022846 ow54a11.s
11	299	19.8	299	1	AI028096 ow51b10.x
12	288	19.1	451	1	AI346780 qp53b04.x
13	283	18.7	522	2	BF432960 7n23a11.x
14	276	18.3	1109	3	CR615693 full-len
15	262	17.3	349	1	AA983880 op63c03.s
16	256	16.9	281	2	AW136718 UI-H-B11-
17	252	16.7	388	2	AW977184 EST389293
18	237	15.7	292	1	AI094420 ou87a01.s
19	230	15.2	630	4	BG770004 602745117
20	223	14.8	350	1	AI357854 qv73h02.x
21	212	14.0	356	1	AA828321 ob55f06.s
22	212	14.0	383	1	AA013045 ze26e05.s
23	212	14.0	392	1	AA279980 zt08e05.s
24	205	13.6	464	1	AV727296 AV727296

25	204	13.5	1056	7	CF552953	CF552953 AGENCOURT
26	203	13.4	254	1	AI242295	AI242295 q116g06.x
27	188	12.4	369	1	AA053949	AA053949 zf48a05.s
28	182	12.0	390	1	AI538997	AI538997 tp76d09.x
29	176	11.6	488	7	R93946	R93946 yt72e09.s1
30	176	11.6	892	5	BQ955258	BQ955258 AGENCOURT
31	171	11.3	791	4	EG741953	EG741953 602635087
32	170	11.3	901	1	AL542834	AL542834 AL542834
33	169	11.2	169	1	AI356664	AI356664 qy16g08.x
34	159	10.5	224	1	AA013207	AA013207 ze26g05.s
35	146	9.7	439	1	AA814860	AA814860 oc06b05.s
36	137	9.1	235	1	AA730023	AA730023 nz89h09.s
37	133	8.8	188	1	AA807164	AA807164 oc36e12.s
38	128	8.5	374	1	AI826340	AI826340 wk44d06.x
39	128	8.5	391	5	BU580066	BU580066 im91d02.y
40	128	8.5	431	2	AW449776	AW449776 UI-H-BI3-
41	128	8.5	466	5	BU734221	BU734221 UI-E-CQ1-
42	128	8.5	484	4	BG482317	BG482317 602526653
43	128	8.5	521	5	BU622650	BU622650 UI-H-FL1-
44	128	8.5	839	6	CD110008	CD110008 AGENCOURT
45	119	7.9	385	5	BU579816	BU579816 im91d02.x

ALIGNMENTS

RESULT 1
AI672308
LOCUS ty63f03.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2283773 3', similar to contains_TAR1.t2 MSRI MSRI repetitive element ; , mRNA sequence.
DEFINITION AI672308.1 GI:4852039
ACCESSION
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 523)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 490.

FEATURES
source

Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2283773"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M.

QY 554 CAGGTCCCTC 563
Db 11 CAGGTCCCTC 2

RESULT 5
AI459114
LOCUS
DEFINITION
tj65c02.x1 Soares NSF F8 9W OT PA P S1 Homo sapiens cDNA clone
IMAGE:2146370 3' similar to contains TAR1.t2 MSRI repetitive
element ;, mRNA sequence.

AI459114 526 bp mRNA linear EST 13-APR-1999
AI459114-1 GI:4311693
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 526)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 653 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 452.
Location/Qualifiers
1. .526
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2146370"
/lab_host="DH10B"
/clone_lib="Soares NSF F8 9W OT PA P S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and cloneIDs: Soares NbHSF pool 1:
309384-310919, 323208-325895 Soares Nb2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NbHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NbHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 25.6%; Score 387; DB 1; Length 526;
Best Local Similarity 99.8%; Pred. No. 9.8e-189;
Matches 437; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1074 CTTCTGCAGCTGCCGCTACGTGAGCCCTTCTCTGGACAGAGCGGCTGGAGAGTATC 1133
Db 1 CTTCTGCAGCTGCCGCTACGTGAGCCCTTCTCTGGACAGAGCGGCTGGAGAGTATC 60

QY 1134 TGTACCCGGCGGGCTGCCGCGCCCGTCCCGTGCCTATACCCCGGCATCCCGCCCGG 1193
Db 61 TGTACCCGGCGGGCTGCCGCGCCCGTCCCGTGCCTATACCCCGGCATCCCGCCCGG 120

QY 1194 CGGCAGCGCGGAGCCGCGCGCTGCCGCGCGCGCGCGCTTCCCTGCTGCTGT 1253
Db 121 CGGCAGCGCGGAGCCGCGCGCTGCCGCGCGCGCGCGCTTCCCTGCTGCTGT 180

QY 1254 CCTCGGTGTGTGCGCCCTCCCGAGAAGCGGGCGCGCGCGACCTCTCTGCGCGC 1313
Db 181 CCTCGGTGTGTGCGCCCTCCCGAGAAGCGGGCGCGCGCGACCTCTCTGCGCGC 240

QY 1314 ACGAGGTGGCGCCCTTGGGGCGCGCACCCCGAGCACCGGCGCGCACCCACCTGC 1373
Db 241 ACGAGGTGGCGCCCTTGGGGCGCGCACCCCGAGCACCGGCGCGCACCCACCTGC 300

QY 1374 CCTTCGCGGGCGCCCGCGAGCGCGGGAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 1433
Db 301 CCTTCGCGGGCGCCCGCGAGCGCGGGAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCGC 360

QY 1434 AGCCAGGAAGAGCTCCCTGAATCCTTGCCTCCCGAAGGACGGAGGTTCAAGCAGAGT 1493
Db 361 AGCCAGGAAGAGCTCCCTGAATCCTTGCCTCCCGAAGGACGGAGGTTCAAGCAGAGT 420

QY 1494 GAGAAAGTAAATACCT 1511
Db 421 GAGAAAGTAAATACCT 438

RESULT 6
BE857082/c
LOCUS
DEFINITION
BE857082 515 bp mRNA linear EST 29-SEP-2000
7923f11.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:3307341 3'
similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND HAIRY-RELATED
PROTEIN 1. ;, mRNA sequence.

ACCESSION BE857082
VERSION BE857082.1 GI:10370753
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 515)
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGA), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 459.
Location/Qualifiers
1. .515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3307341"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn23"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCATATCTTTTTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

```
Query Match      25.1%; Score 380; DB 2; Length 515;
Best Local Similarity 99.8%; Pred. No. 4.2e-185;
Matches 430; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAAGAGACCGAATTAATGAA 193
Db 431 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAAGAGACCGAATTAATGAA 372

QY 194 TGCATTGCTCAGCTGAAAGATTACTGCGCTGAACATCTGAAATTGACAACCTCTGGGACAT 253
Db 371 TGCATTGCTCAGCTGAAAGATTACTGCGCTGAACATCTGAAATTGACAACCTCTGGGACAT 312

QY 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAAACACTTAAAGCTTTAACCGCCTTA 313
Db 311 CTGGAAAAGCTGTAGTCTTGGAAATTAACCTTTGAAAACACTTAAAGCTTTAACCGCCTTA 252

QY 314 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 373
Db 251 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 192

QY 374 CCCATTGCTCCGACTTGGATGCGTTCCACTCGGATTTCAAACATGCGCCAAAGAGTC 433
Db 191 CCCATTGCTCCGACTTGGATGCGTTCCACTCGGATTTCAAACATGCGCCAAAGAGTC 132

QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCCGGTGTGTCCAGCTG 493
Db 131 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCCGGTGTGTCCAGCTG 72

QY 494 ATCAACCACTTGACCGCCGTGGCCACCCAGTCTTGTCCACCCAGGAGCTGTGACTCAA 553
Db 71 ATCAACCACTTGACCGCCGTGGCCACCCAGTCTTGTCCACCCAGGAGCTGTGACTCAA 12

QY 554 CAGTCCCTCT 564
Db 11 CAGTCCCTCT 1
```

```
RESULT 7
AW073334/c
LOCUS
DEFINITION
  wy97905.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:2556536 3'
  similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND HAIRY-RELATED
  PROTEIN 1. ; mRNA sequence.
AW073334
AW073334.1 GI:6028332
EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 557)
  NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute / National Institute of Neurological
  Disorders and Stroke, Brain Tumor Genome Anatomy Project
  (CGAP/BTGA), Tumor Gene Index
  Unpublished (1998)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
  Ph.D.
  CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
  Bonaldo, Ph.D.
  CDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Seq primer: -40UP from Gibco
  High quality sequence stop: 455.
  Location/Qualifiers
    1..557
    /organism="Homo sapiens"
    /mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:2556536"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn23"
/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATATCTTTTCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
```

```
ORIGIN
Query Match      25.1%; Score 380; DB 2; Length 557;
Best Local Similarity 99.8%; Pred. No. 4.2e-185;
Matches 430; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAAGAGACCGAATTAATGAA 193
Db 431 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAAGAGACCGAATTAATGAA 372

QY 194 TGCATTGCTCAGCTGAAAGATTACTGCGCTGAACATCTGAAATTGACAACCTCTGGGACAT 253
Db 371 TGCATTGCTCAGCTGAAAGATTACTGCGCTGAACATCTGAAATTGACAACCTCTGGGACAT 312

QY 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAAACACTTAAAGCTTTAACCGCCTTA 313
Db 311 CTGGAAAAGCTGTAGTCTTGGAAATTAACCTTTGAAAACACTTAAAGCTTTAACCGCCTTA 252

QY 314 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 373
Db 251 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 192

QY 374 CCCATTGCTCCGACTTGGATGCGTTCCACTCGGATTTCAAACATGCGCCAAAGAGTC 433
Db 191 CCCATTGCTCCGACTTGGATGCGTTCCACTCGGATTTCAAACATGCGCCAAAGAGTC 132

QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCCGGTGTGTCCAGCTG 493
Db 131 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCCGGTGTGTCCAGCTG 72

QY 494 ATCAACCACTTGACCGCCGTGGCCACCCAGTCTTGTCCACCCAGGAGCTGTGACTCAA 553
Db 71 ATCAACCACTTGACCGCCGTGGCCACCCAGTCTTGTCCACCCAGGAGCTGTGACTCAA 12

QY 554 CAGTCCCTCT 564
Db 11 CAGTCCCTCT 1
```

```
RESULT 8
AA996006/c
LOCUS
DEFINITION
  os26c04.s1 NCI CGAP Kid5 Homo sapiens cDNA clone IMAGE:1606470 3'
  similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND HAIRY-RELATED
  PROTEIN 1. ; mRNA sequence.
AA996006
AA996006.1 GI:3182495
EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 434)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
```


Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 496 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 291.

FEATURES
source
1. 434
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1606470"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid5"
/note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAAATCGCGCCGCAATATTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 21.3%; Score 322; DB 1; Length 434;
Best Local Similarity 100.0%; Pred. No. 4.8e-155;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 243 CTCCTGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTGAACACACTTAAAGCTT 302
Db |||||||
QY 322 CTCCTGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTGAACACACTTAAAGCTT 263
Db |||||||
QY 303 TAACCGCCTTAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATGGGAGCGAT 362
Db |||||||
QY 262 TAACCGCCTTAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATGGGAGCGAT 203
Db |||||||
QY 363 CTCCTGAATCGCCCATTCAGTCGACTTGGATGCGTTCCACTCGGGATTTCAACATGCG 422
Db |||||||
QY 202 CTCCTGAATCGCCCATTCAGTCGACTTGGATGCGTTCCACTCGGGATTTCAACATGCG 143
Db |||||||
QY 423 CCAAGAACTCTTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGAGCGCGGT 482
Db |||||||
QY 142 CCAAGAACTCTTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGAGCGCGGT 83
Db |||||||
QY 483 GTGTCCAGTGATCAACCACTTGACAGCCCGTGGCCACCAGTCTTGTCCACCCGCGAGC 542
Db |||||||
QY 82 GTGTCCAGTGATCAACCACTTGACAGCCCGTGGCCACCAGTCTTGTCCACCCGCGAGC 23
Db |||||||
QY 543 TGTGACTCAACAGGTCCCTCT 564
Db |||||||
QY 22 TGTGACTCAACAGGTCCCTCT 1
Db |||||||

RESULT 9
CG464925
LOCUS
DEFINITION
CG464925 310 bp DNA linear GSS 24-SEP-2003
KRIBB_2D_D67T7 pBS-NE Library Homo sapiens genomic clone
KRIBB_2D_D67 5, genomic survey sequence.
ACCESSION
CG464925 GI:35188360
VERSION
GSS.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 310)

AUTHORS
TITLE
JOURNAL
COMMENT
Lee, K.T., Kim, J.H., Hahn, Y.S., Yang, J.O., Chu, M.Y., Kim, H.C.,
Oh, K.J., Kim, S.S., Yoo, H.S. and Kim, Y.S.
Rapid isolation of NotI-linked CpG island sequences from RLGS gel
Unpublished (2003)
Contact: KIM YS
Human Genome Research Laboratory
Korea Research Institute of Bioscience and Biotechnology
52 Eoeun-dong, Yuseong-gu, Daejeon 305-333, Republic of Korea
Tel: 82-42-879-8110
Fax: 82-42-879-8119
Email: yongsung@kribb.re.kr
This result was produced by the program for Functional Analysis of
Human Genome funded from Ministry of Science and Technology,
Republic of Korea. Clones are available from the Center for
Functional Analysis of Human Genome. See URL:
http://21cgenome.kribb.re.kr/ for details or contact:
yongsung@kribb.re.kr.
Seq primer: T7 primer
Class: NotI site
High quality sequence stop: 310.

FEATURES

source
Location/Qualifiers
1. 310
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="KRIBB_2D_D67"
/sex="male"
/tissue_type="blood"
/dev_stage="adult"
/clone_lib="pBS-NE Library"
/note="Vector: pBluescriptKS(+); Site 1: NotI; Site 2:
EcoRV; Genomic DNA was digested with NotI and EcoRV and
the resulting fragments were ligated into ZAP1/NotI
vector DNA. After the digestion with EcoRV, the mixture
was run on 0.8% LMP agarose gel, eluted a gel portion over
22 kb and purified with beta-agarase I. The purified DNA
was digested with NotI, run on 1% LMP agarose gel, and
eluted with Gel Extraction kit the DNA fragments ranging
0.7 to 4 kb, which cover on first-dimensional gel
comprising the central portion of the standard RLGS
profile. The DNA solution was ligated into pBluescript
KS(+) with NotI and EcoRV ends. The ligated solution was
transformed into DH5a cell using electroporation machine. "

ORIGIN

Query Match 20.5%; Score 310; DB 9; Length 310;
Best Local Similarity 100.0%; Pred. No. 7.9e-149;
Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1028 GCGGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTGCTTCCCTCTCGCCTTCTGCAGTGCC 1087
Db |||||||
QY 1 1 GCGGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTGCTTCCCTCTCGCCTTCTGCAGTGCC 60
Db |||||||
QY 1088 GCCTACGTGCAGCCCTTCTCTGGACAAGAGCGGCGCTGGAGAAGTATCTGTACCCGCGCGCG 1147
Db |||||||
QY 61 GCCTACGTGCAGCCCTTCTCTGGACAAGAGCGGCGCTGGAGAAGTATCTGTACCCGCGCGCG 120
Db |||||||
QY 1148 GCTGCCGCGCGCGCGCTTCCGCTGCTATACCCCGGCATCCCGCGCGCGCGCGCGCGCA 1207
Db |||||||
QY 121 GCTGCCGCGCGCGCTTCCGCTGCTATACCCCGGCATCCCGCGCGCGCGCGCGCGCA 180
Db |||||||
QY 1208 GCCGCGCGCGCGCGCTGCCGCGCGCGCGCGCGCTTCCCGCTGCTCTCGGTGTGTCG 1267
Db |||||||
QY 181 GCCGCGCGCGCGCGCTGCCGCGCGCGCGCGCGCTTCCCGCTGCTCTCGGTGTGTCG 240
Db |||||||
QY 1268 CCCCTCCCGAGAGCGGCG 1327
Db |||||||
QY 241 CCCCTCCCGAGAGCGGCG 300
Db |||||||
QY 1328 CTTGGGCGCG 1337
Db |||||||
QY 301 CTTGGGCGCG 310

```
RESULT 10
AI022846/c
LOCUS
DEFINITION
  AI022846
  ow54a11.s1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone
  IMAGE:1650620 3' similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND
  HAIRY-RELATED PROTEIN 1. ;, mRNA sequence.
ACCESSION
  AI022846
VERSION
  AI022846.1 GI:3238087
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
  1 (bases 1 to 301)
REFERENCE
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished (1997)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
  Bonaldo, Ph.D.
  CDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 317 Std Error: 0.00
  Seq primer: -40m13 fwd. ET from Amersham.
FEATURES
  Location/Qualifiers
    1..301
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:1650620"
      /tissue_type="parathyroid tumor"
      /dev_stage="adult"
      /lab_host="DH10B (ampicillin resistant)"
      /clone_lib="Soares parathyroid tumor NbHPA"
      /note="Organ: parathyroid gland; Vector: pT7T3D
      (Pharmacia) with a modified polylinker; Site 1: Not I;
      Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
      oligo(dT) primer
      [5'-TGTTACCAATCTGAAGTGGAGCGCGCCACCAATTTTTTTTTTTTTTTT
      TTTTT-3'], double-stranded cDNA was size selected, ligated
      to Eco RI adapters (Pharmacia), digested with Not I and
      cloned into the Not I and Eco RI sites of a modified pT7T3
      vector (Pharmacia). Library went through one round of
      normalization to a Cot = 5. Library constructed by Bento
      Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
      adenomas was kindly provided by Dr. Stephen Marx, National
      Institute of Diabetes and Digestive and Kidney Diseases,
      NIH."
ORIGIN
  Query Match
  Best Local Similarity 19.9%; Score 301; DB 1; Length 301;
  Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 267 TAGTCTTGGAAATTAACCTTTGAACACACTTAAAGCTTTAACCGCCTTAACCGAGCAACAGC 326
  Db
  301 TAGTCTTGGAAATTAACCTTTGAACACACTTAAAGCTTTAACCGCCTTAACCGAGCAACAGC 242
  QY 327 ATCAGAAGATATATGCTTTACAGAAATGGGAGCGGATCTCTGAAATGCGCCCATTCAGTCCG 386
  Db
  241 ATCAGAAGATATATGCTTTACAGAAATGGGAGCGGATCTCTGAAATGCGCCCATTCAGTCCG 182
  QY 387 ACTTGGATGCGGTCCACTCGGGATTTCAACATCGGCCAAAGAGTCTTGCAATACCTCT 446
  Db
  181 ACTTGGATGCGGTCCACTCGGGATTTCAACATCGGCCAAAGAGTCTTGCAATACCTCT 122
  QY 447 CCCGGTTGAGAGTGGAACACCCAGGAGCGCGGTGTGTCCAGTGTATCAACCACTTGC 506
```

```
Db 121 CCCGGTTTGAGAGCTGACACCCAGGAGCGCGGTGTGTCCAGTGTATCAACCACTTGC 62
QY 507 ACGCCGTGGCCACCCAGTCTTGTGCCACCCCGAGCTGTTGACTCAACAGGTCCTCTGA 566
Db 61 ACGCCGTGGCCACCCAGTCTTGTGCCACCCCGAGCTGTTGACTCAACAGGTCCTCTGA 2
QY 567 G 567
Db 1 G 1
RESULT 11
AI028096/c
LOCUS
DEFINITION
  AI028096
  ow51b10.x1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone
  IMAGE:1650331 3' similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND
  HAIRY-RELATED PROTEIN 1. ;, mRNA sequence.
ACCESSION
  AI028096
VERSION
  AI028096.1 GI:3245405
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
  1 (bases 1 to 299)
REFERENCE
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished (1997)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
  Bonaldo, Ph.D.
  CDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 331 Std Error: 0.00
  Seq primer: -40m13 fwd. ET from Amersham.
FEATURES
  Location/Qualifiers
    1..299
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:1650331"
      /tissue_type="parathyroid tumor"
      /dev_stage="adult"
      /lab_host="DH10B (ampicillin resistant)"
      /clone_lib="Soares parathyroid tumor NbHPA"
      /note="Organ: parathyroid gland; Vector: pT7T3D
      (Pharmacia) with a modified polylinker; Site 1: Not I;
      Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
      oligo(dT) primer
      [5'-TGTTACCAATCTGAAGTGGAGCGCGCCACCAATTTTTTTTTTTTTTTT
      TTTTT-3'], double-stranded cDNA was size selected, ligated
      to Eco RI adapters (Pharmacia), digested with Not I and
      cloned into the Not I and Eco RI sites of a modified pT7T3
      vector (Pharmacia). Library went through one round of
      normalization to a Cot = 5. Library constructed by Bento
      Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
      adenomas was kindly provided by Dr. Stephen Marx, National
      Institute of Diabetes and Digestive and Kidney Diseases,
      NIH."
ORIGIN
  Query Match
  Best Local Similarity 19.8%; Score 299; DB 1; Length 299;
  Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 266 GTAGTCTTGGAAATTAACCTTTGAACACACTTAAAGCTTTAACCGCCTTAACCGAGCAACAG 325
```

Db 299 GTAGCTTGAATTAACCTTTGAACACTTAAAGCTTTAACCGCCTTAACCGCAACAG 240
QY 326 CATCAGACATAATTGCTTTACAGATGGGAGCGATCTCTGAAATCGCCCATTCAGTCC 385
Db 239 CATCAGACATAATTGCTTTACAGATGGGAGCGATCTCTGAAATCGCCCATTCAGTCC 180
QY 386 GACTTGGATCGGTTCCACTCGGATTTCAAAACATCGGCCAAAGAGTCTTGCAATACCTC 445
Db 179 GACTTGGATCGGTTCCACTCGGATTTCAAAACATCGGCCAAAGAGTCTTGCAATACCTC 120
QY 446 TCCCGGTTGAGAGCTGGACACCCAGGAGCCGCGGTGTGTCCAGCTGATCAACCACTTG 505
Db 119 TCCCGGTTGAGAGCTGGACACCCAGGAGCCGCGGTGTGTCCAGCTGATCAACCACTTG 60
QY 506 CACGCCGTGGCCACCCAGTTCTTGCCACCCCGCAGCTGTGACTCAACAGTCCCTCT 564
Db 59 CACGCCGTGGCCACCCAGTTCTTGCCACCCCGCAGCTGTGACTCAACAGTCCCTCT 1

RESULT 12
AI346780 451 bp mRNA linear EST 02-FEB-1999
LOCUS qp53b04.x1 NCI CGAP Co8 Homo sapiens cDNA clone IMAGE:1926703 3'
DEFINITION similar to contains TAR1.b1 MSRI repetitive element ;, mRNA
sequence.
ACCESSION AI346780
VERSION AI346780.1 GI:4083986
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 451)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 2359 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1926703"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Co8"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

FEATURES
source
1. .451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1926703"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Co8"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 19.1%; Score 288; DB 1; Length 451;
Best Local Similarity 100.0%; Pred. No. 2e-137;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1224 CCGCCGCGCGCGGTTCCCTGCTGCTCGGTGTGTGCGCCCTCCCGAGAAGG 1283

Db 158 CCGCCGCGCGCGGTTCCCTGCTGCTCGGTGTGTGCGCCCTCCCGAGAAGG 217
QY 1284 CGGGCGCGCGCGCGACCTCTCTCGCGCACGAGTGGCGCCCTTGGGGCGCGCAC 1343
Db 218 CGGGCGCGCGCGCGACCTCTCTCGCGCACGAGGTGGCGCCCTTGGGGCGCGCAC 277
QY 1344 CCCAGCACCCCGCACGGCCGCGACCCACCTGCCCTTCGCCGGGCCCCCGGGAACC 1403
Db 278 CCCAGCACCCCGCACGGCCGCGACCCACCTGCCCTTCGCCGGGCCCCCGGGAACC 337
QY 1404 CGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGAAAGCTCCCTGAATCCTTG 1463
Db 338 CGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGAAAGCTCCCTGAATCCTTG 397
QY 1464 CGTCCCGAAGGACGGAGGTTCAAGCAGAGTGAGAAAGTAAAAATACCCT 1511
Db 398 CGTCCCGAAGGACGGAGGTTCAAGCAGAGTGAGAAAGTAAAAATACCCT 445

RESULT 13
BF432960 522 bp mRNA linear EST 29-NOV-2000
LOCUS 7n23a11.x1 NCI CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3565341 3'
DEFINITION similar to contains element TAR1 repetitive element ;, mRNA
sequence.
ACCESSION BF432960
VERSION BF432960.1 GI:11445123
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 522)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 484.
Location/Qualifiers
1. .522
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3565341"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu24"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI CGAP_Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (clones
1414920-1417991 and 1520904-152439). Subtraction by Bento
Soares and M. Fatima Bonaldo."

FEATURES
source
1. .522
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3565341"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu24"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI CGAP_Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (clones
1414920-1417991 and 1520904-152439). Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 18.7%; Score 283; DB 2; Length 522;
Best Local Similarity 100.0%; Pred. No. 8e-135;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
QY 1229 GCCGCCGCCGCTTCCCTGCCTGTCTCGTGTGTTGTGCGCCCTCCCGAGAGCGGGC 1288
Db 5 GCCGCCGCCGCTTCCCTGCCTGTCTCGTGTGTTGTGCGCCCTCCCGAGAGCGGGC 64

QY 1289 GCCGCCGCCGCTTCCCTGCCTGTCTCGGACGAGGTGGCGCCCTTGGGCGCCGCAACCCAG 1348
Db 65 GCCGCCGCCGCTTCCCTGTCTCGCGACGAGGTGGCGCCCTTGGGCGCCGCAACCCAG 124

QY 1349 CACCCGCACGGCGCACCCACCTGCCCTTCGCCGGCCCGCCCGAGCGGGGAACCCGAG 1408
Db 125 CACCCGCACGGCGCACCCACCTGCCCTTCGCCGGCCCGCCCGAGCGGGGAACCCGAG 184

QY 1409 AGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGGAGCTCCCTGAATCCTTGGCTCC 1468
Db 185 AGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGGAGCTCCCTGAATCCTTGGCTCC 244

QY 1469 CGAAGGACGGAGTTCAAGCAGACTGAGAGTTAAATACCT 1511
Db 245 CGAAGGACGGAGTTCAAGCAGACTGAGAGTTAAATACCT 287

RESULT 14
CR615693
LOCUS
DEFINITION
full-length cDNA clone CS0DE013YK10 of Placenta of Homo sapiens
(human).
CR615693
ACCESSION
CR615693.1 GI:50496500
VERSION
HTC; CNSLT_cDNA.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1109)
AUTHORS
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
REFERENCE
2 (bases 1 to 1109)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (20-JUL-2004) Genoscope - Centre National de Sequenage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1. .1109
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013YK10"
/tissue type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 18.3%; Score 276; DB 3; Length 1109;
Best Local Similarity 100.0%; Pred. No. 3.5e-131;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1236 CCGCGTTCCTGCTGCTCGGTGTTGTCGCCCTCCCGAGAGCGGGCGCGCG 1295
Db 1 CCGCGTTCCTGCTGCTCGGTGTTGTCGCCCTCCCGAGAGCGGGCGCGCGCG 60

QY 1296 CCGCGACCTCTCGCGACGAGGTGGCGCCCTTGGGGCGCGACCCCGACCCGC 1355
Db 61 CCGCGACCTCTCTGCGCACGAGGTGGCGCCCTTGGGGCGCGACCCCGACCCGC 120
```

```
QY 1356 ACGGCCGACCCACCTGCTCCCTTCGCCGGGCCCGGAGCCGGGAACCGGAGAGCTCTG 1415
Db 121 ACGGCCGACCCACCTGCTCCCTTCGCCGGGCCCGGAGCCGGGAACCGGAGAGCTCTG 180

QY 1416 CTCAGGAAGATCCCTCGCAGCCAGGAAGGAGCTCCCTGAATCCTTGGTCCCGAAGGA 1475
Db 181 CTCAGGAAGATCCCTCGCAGCCAGGAAGGAGCTCCCTGAATCCTTGGTCCCGAAGGA 240

QY 1476 CGGAGTTCAAGCAGAGTGAGAGTTAAATACCT 1511
Db 241 CGGAGTTCAAGCAGAGTGAGAGTTAAATACCT 276

RESULT 15
AA983880/c
LOCUS
DEFINITION
op63c03.s1 Soares_NFL T GBC_S1 Homo sapiens cDNA clone
IMAGE:1581508 3' similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND
HAIRY-RELATED PROTEIN 1. ; mRNA sequence.
AA983880
ACCESSION
AA983880.1 GI:3162405
VERSION
EST.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 349)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 612 Std Error: 0.00
Seg primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 348.
FEATURES
Location/Qualifiers
source
1. .349
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1581508"
/lab_host="DH10B"
/clone_lib="Soares_NFL T GBC_S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBH19W, testis NHT, and B-cell
NCI-CGAP_GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
```

```
ORIGIN
Query Match 17.3%; Score 262; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 6.1e-124;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGAATTAACCTTGAACACTTAAAGCT 301
Db 322 ACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGAATTAACCTTGAACACTTAAAGCT 263

QY 302 TTAACCGCCTTAACCGAGCAACAGCATCAGAGATAATTGCTTTACAGATGGGAGCGA 361
Db 262 TTAACCGCCTTAACCGAGCAACAGCATCAGAGATAATTGCTTTACAGATGGGAGCGA 203
```

QY	362	TCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGGATTTCAAACATGC	421
Db	202	TCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGGATTTCAAACATGC	143
QY	422	GCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGCGG	481
Db	142	GCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGCGG	83
QY	482	TGTGTCCAGCTGATCAACCACT	503
Db	82	TGTGTCCAGCTGATCAACCACT	61

Search completed: December 17, 2004, 16:09:38
Job time : 4567.79 secs


```
QY 1423 CCGCGACCCCTCCTGCGCACGAGGTGGCGCCCTTGGGGCGCCGACACCCCGCACCCCGC 1482
Db 61 CCGCGACCCCTCCTGCGCACGAGGTGGCGCCCTTGGGGCGCCGACACCCCGCACCCCGC 120
QY 1483 ACGGCCGACCCACCTGCGCCCTTCCGCGGGGCGCGAGCCGGGGAACCCGGAGAGCTCTG 1542
Db 121 ACGGCCGACCCACCTGCGCCCTTCCGCGGGGCGCGAGCCGGGGAACCCGGAGAGCTCTG 180
QY 1543 CTCAGGAAGATCCCTCGCAGCCAGGAAAGGAAGCTCCCTGAATCCCTGCGTCCCGAAGGA 1602
Db 181 CTCAGGAAGATCCCTCGCAGCCAGGAAAGGAAGCTCCCTGAATCCCTGCGTCCCGAAGGA 240
QY 1603 CGGAGGTTCAAGCAGAGTGAGAAATTAACCTTAAGGAGGTTCAAGCAGAGTGAGA 1662
Db 241 CGGAGGTTCAAGCAGAGTGAGAAATTAACCTTAAGGAGGTTCAAGCAGAGTGAGA 300
QY 1663 AGTTAAATACCTTAAGGTCTTTTAAGGGAAGAGTGAATAGATGCACGACAGGCATAA 1722
Db 301 AGTTAAATACCTTAAGGTCTTTTAAGGGAAGAGTGAATAGATGCACGACAGGCATAA 360
QY 1723 ACAAGAACAAACAACAGGTGTTATGTGTACATTCGGAGTTCCTGTTTGTCTCATCCCGC 1782
Db 361 ACAAGAACAAACAACAGGTGTTATGTGTACATTCGGAGTTCCTGTTTGTCTCATCCCGC 420
QY 1783 ACCACCCCACTCCACACACTAATCCCTTCTTCCCGCCACAGCTGTAAGATCC 1842
Db 421 ACCACCCCACTCCACACACTAATCCCTTCTTCCCGCCACAGCTGTAAGATCC 480
QY 1843 TATGCGAAAGACACTGGCTCTTTTATATCCCGCCCAATAAATTTGCCCCCTTTTAGG 1902
Db 481 TATGCGAAAGACACTGGCTCTTTTATATCCCGCCCAATAAATTTGCCCCCTTTTAGG 540
QY 1903 CCATGTTCCATTATCTTAAATTTGGAACCTAATTCGAGAGGAAGTAAGAGGTTCTGT 1962
Db 541 CCATGTTCCATTATCTTAAATTTGGAACCTAATTCGAGAGGAAGTAAGAGGTTCTGT 600
QY 1963 TCTGTGGCTGAGCTAGGTGAACCCCGGGGAGGGAAGATGTTAACAACCTTTGACGTCT 2022
Db 601 TCTGTGGCTGAGCTAGGTGAACCCCGGGGAGGGAAGATGTTAACAACCTTTGACGTCT 660
QY 2023 TTGGAGTTGACATGGAACAGCAGGTAGTTGTTATGTAGAGCTAGTCTCAAGCTGCCCT 2082
Db 661 TTGGAGTTGACATGGAACAGCAGGTAGTTGTTATGTAGAGCTAGTCTCAAGCTGCCCT 720
QY 2083 GCCTGTTTATAGGCGCTTCCACAAACAGATTGAGGCTCTTTTATGAATTTTACTC 2142
Db 721 GCCTGTTTATAGGCGCTTCCACAAACAGATTGAGGCTC-TTTTATGAATTTTACTC 779
QY 2143 TTCAGTATTTTCTAATGTTTCAAGCTTTCTAAGGCAATATATTTTCAAGAAAGTGAGGAT 2202
Db 780 TTCAGTATTTTCTAATGTTTCAAGCTTTCTAAGGCAATATATTTTCAAGAAAGTGAGGAT 839
QY 2203 GCAGTTTCTCAGCTGCAACCTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAACTTG 2262
Db 840 GCAGTTTCTCAGCTGCAACCTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAACTTG 899
QY 2263 CACTCGTTAAAGAAACACGGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGG 2322
Db 900 CACTCGTTAAAGAAACACGGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGG 959
QY 2323 ATGAGAAGGCCAGAAATCATTCTAGTACATTTGCTAACAATTTATGAGAAATTGACCAT 2382
Db 960 ATGAGAAGGCCAGAAATCATTCTAGTACATTTGCTAACAATTTATGAGAAATTGACCAT 1019
QY 2383 GAATTAATGGACTCATCTTAATTTCTCTAAGTCCATATATAGATAGATATCTATCTGTA 2442
Db 1020 GAATTAATGGACTCATCTTAATTTCTCTAAGTCCATATATAGATAGATATCTATCTGTA 1079
QY 2443 CAGATTTCTATTATCCATAGATAGGTATC 2472
Db 1080 CAGATTTCTATTATCCATAGATAGGTATC 1109
```

```
RESULT 2
BM542244
LOCUS
DEFINITION AGENCOURT_6436506 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539636
5', mRNA sequence.
ACCESSION BM542244
VERSION BM542244.1 GI:18771584
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 715)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1234 row: h column: 05
High quality sequence stop: 635.
FEATURES
Location/Qualifiers
1..715
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5539636"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
```

ORIGIN

```
Query Match 18.8%; Score 683; DB 4; Length 715;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 AGGGAGAGCGAGAGAGAGACTGGAGACGCGACAGATCCCCCAAGGTCTCCCAAGCCTACC 71
Db 3 AGGGAGAGCGAGAGAGAGACTGGAGACGCGACAGATCCCCCAAGGTCTCCCAAGCCTACC 62
QY 72 GTCCCAACAGATTATTGTACAGAGCCCCCAAAATCGAAACAGAGGAAACGACAGCAGTTG 131
Db 63 GTCCCAACAGATTATTGTACAGAGCCCCCAAAATCGAAACAGAGGAAACGACAGCAGTTG 122
QY 132 AACATGGACGAAGAAATTCCTCATTTTGAAGAGAGACAGTTACTGGAACATAGAGATTTT 191
Db 123 AACATGGACGAAGAAATTCCTCATTTTGAAGAGAGACAGTTACTGGAACATAGAGATTTT 182
QY 192 ATAGGACTGGACTATTCTCTTTGTATATGTGTAAACCCAAAGAGAGCATGAAACGAGAC 251
Db 183 ATAGGACTGGACTATTCTCTTTGTATATGTGTAAACCCAAAGAGAGCATGAAACGAGAC 242
QY 252 GACACCAAGGATACCTACAAATTCGCGACAGATTAAATAGAAAGAAAGAGAGACCGA 311
Db 243 GACACCAAGGATACCTACAAATTCGCGACAGATTAAATAGAAAGAAAGAGAGACCGA 302
QY 312 ATTAATGAATGCATTCGTCAGCTGAAGATTTACTGCTGAAATTTGAAATTTGACAACT 371
Db 303 ATTAATGAATGCATTCGTCAGCTGAAGATTTACTGCTGAAATTTGAAATTTGACAACT 362
QY 372 CTGGGACATCTGGAGAAAGCTGTAGTCTTTGGAATTAACCTTTGAAACACTTAAAGCTTTA 431
```

Db 363 CTGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTA 422

QY 432 ACCGCCTTAACCGAGCAACAGCATCAGAAGATAATGCTTTACAGAATGGGAGCGATCT 491

Db 423 ACCGCCTTAACCGAGCAACAGCATCAGAAGATAATGCTTTACAGAATGGGAGCGATCT 482

QY 492 CTGAAATCGCCCATTCAGTCCGACTTGGATCGGTTCCACTCGGGATTTCAAACATGCGCC 551

Db 483 CTGAAATCGCCCATTCAGTCCGACTTGGATCGGTTCCACTCGGGATTTCAAACATGCGCC 542

QY 552 AAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGGTGT 611

Db 543 AAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGGTGT 602

QY 612 GTCCAGCTGATCAACCACTTGACGCCGTCGTCGCCACCCAGTTCTTGCCACCCCGCAGCTG 671

Db 603 GTCCAGCTGATCAACCACTTGACGCCGTCGTCGCCACCCAGTTCTTGCCACCCCGCAGCTG 662

QY 672 TTGACTCAACAGGTCCTCTGAG 694

Db 663 TTGACTCAACAGGTCCTCTGAG 685

RESULT 3

BQ955258 892 bp mRNA linear EST 21-AUG-2002

LOCUS AGENCOURT_8732698 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6455656

DEFINITION 5', mRNA sequence.

ACCESSION BQ955258

VERSION BQ955258.1 GI:22370736

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 892)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ARCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Plate: LLCM2634 row: k column: 17 High quality sequence stop: 684.

FEATURES

source

1. .892

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6455656"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 101"

/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

ORIGIN

Query Match 18.2%; Score 663; DB 5; Length 892;

Best Local Similarity 100.0%; Pred. No. 2.9e-305;

Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1463 GCCGACCCCCAGCACCCGCGCGCGCACCCACCTGCCCCCTTCGCCGGCCCCCGGAGCC 1522

Db 48 GCCGACCCCCAGCACCCGCGCGCGCACCCACCTGCCCCCTTCGCCGGCCCCCGGAGCC 107

QY 1523 GGGAAACCCGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAAGGAGTCCCTG 1582

Db 108 GGGAAACCCGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAAGGAGTCCCTG 167

QY 1583 AATCCTTGCGTCCCGAAGGAGCGGAGTTCAAGCAGAGTGAGAAGTTAAAAATACCTTAAG 1642

Db 168 AATCCTTGCGTCCCGAAGGAGCGGAGTTCAAGCAGAGTGAGAAGTTAAAAATACCTTAAG 227

QY 1643 GAGGTTCAAGCAGAGTGAGAAGTTAAAAATACCTTAAGTCTTTAAGGAGGAAAGTGTA 1702

Db 228 GAGGTTCAAGCAGAGTGAGAAGTTAAAAATACCTTAAGTCTTTAAGGAGGAAAGTGTA 287

QY 1703 TAGATGCACGACAGGCGATAAACAAGAAACAACAACAGAGTGTATGTGTACATTCCGAGT 1762

Db 288 TAGATGCACGACAGGCGATAAACAAGAAACAACAACAGAGTGTATGTGTACATTCCGAGT 347

QY 1763 TCCTGTTTGTCTCATCCCGGACCCACCCCTCCACACTAATCCCTTCTTCCCC 1822

Db 348 TCCTGTTTGTCTCATCCCGGACCCACCCCTCCACACTAATCCCTTCTTCCCC 407

QY 1823 CCACAGCTGTAAAGATCCTATGCGAAAGACACTGGCTCTTTTAAATCCCCCAAT 1882

Db 408 CCACAGCTGTAAAGATCCTATGCGAAAGACACTGGCTCTTTTAAATCCCCCAAT 467

QY 1883 AAATTTGCCCCCTTTTGGCCATGTTCCATTATCTCTTAAATTTGGAACCTAATTCGAG 1942

Db 468 AAATTTGCCCCCTTTTGGCCATGTTCCATTATCTCTTAAATTTGGAACCTAATTCGAG 527

QY 1943 AGGAAGTAAGAAGGCTCTGTCTGTGGCTGAGCTAGTGAAACCCCGGGTAGGGAAAGA 2002

Db 528 AGGAAGTAAGAAGGCTCTGTCTGTGGCTGAGCTAGTGAAACCCCGGGTAGGGAAAGA 587

QY 2003 TGTAAACACCTTTGACGCTCTTTGGAGTTGACATGGAACAGCAGGTAGTTGTTATGTAGAG 2062

Db 588 TGTAAACACCTTTGACGCTCTTTGGAGTTGACATGGAACAGCAGGTAGTTGTTATGTAGAG 647

QY 2063 CTAGTTCTCAAAGCTGCCCTGCTGTTTGGAGGCGTTCCACAAACAGATTGAGGCTCT 2122

Db 648 CTAGTTCTCAAAGCTGCCCTGCTGTTTGGAGGCGTTCCACAAACAGATTGAGGCTCT 707

QY 2123 TTT 2125

Db 708 TTT 710

RESULT 4

AL542834

LOCUS

DEFINITION

AL542834 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YK10 5-PRIME, mRNA sequence.

ACCESSION

AL542834

VERSION

AL542834.3 GI:45718407

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 901)

AUTHORS

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

On Feb 15, 2001 this sequence version replaced gi:30548385.

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

AL542834 901 bp mRNA linear EST 24-MAR-2004

AL542834 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YK10 5-PRIME, mRNA sequence.

AL542834

AL542834.3 GI:45718407

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 901)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 15, 2001 this sequence version replaced gi:30548385.

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 6711.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0DE013BF05QPI&c=6711.r.

FEATURES		Location/Qualifiers	
source	1..901	/organism="Homo sapiens"	
		/mol_type="mRNA"	
REFERENCE	/db_xref="taxon:9606"		
	/clone="CS0DE013YK10"		
AUTHORS	/tissue type="PLACENTA"		
	/clone lib="Homo sapiens PLACENTA"		
COMMENT	/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."		
ORIGIN		Query Match	
		Best Local Similarity	18.0%; Score 657; DB 1; Length 901;
		Matches 657; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1469	CCCCCAGCACCCGCGCGCACCCACCTGCGCTTCCGCGCGCGCGCGCGCGGAA	1528
Db	106	CCCCCAGCACCCGCGCGCACCCACCTGCGCTTCCGCGCGCGCGCGCGGAA	165
QY	1529	CCCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAAGGAGCTCCCTGAATCCT	1588
Db	166	CCCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAAGGAGCTCCCTGAATCCT	225
QY	1589	TGCGTCCCGAAGGACGAGGTTCAAGCAGAGTGAGAAAGTTAAATACCTTAAAGGAGTT	1648
Db	226	TGCGTCCCGAAGGACGAGGTTCAAGCAGAGTGAGAAAGTTAAATACCTTAAAGGAGTT	285
QY	1649	CAAGCAGAGTGAGAAGTTAAATACCTTAAAGGTTCTTAAAGGAGGAGTGTATAGATG	1708
Db	286	CAAGCAGAGTGAGAAGTTAAATACCTTAAAGGTTCTTAAAGGAGGAGTGTATAGATG	345
QY	1709	CACGACAGGCATAAACAAGAACAACAACACAGGTGTATGTGTACATTCGGAGTTCCTGT	1768
Db	346	CACGACAGGCATAAACAAGAACAACAACACAGGTGTATGTGTACATTCGGAGTTCCTGT	405
QY	1769	TTTGCTCATCCCGACCCACCCCTCCACACACTAACATCCCTTCTTCCCCACCA	1828
Db	406	TTTGCTCATCCCGACCCACCCCTCCACACACTAACATCCCTTCTTCCCCACCA	465
QY	1829	GCTGTAAAGATCCTATCGGAAAGACACATGGCTCTTTTAAATCCCTTAAATTT	1888
Db	466	GCTGTAAAGATCCTATCGGAAAGACACATGGCTCTTTTAAATCCCTTAAATTT	525
QY	1889	TGCCCCCTTTTAGCCATGTTCCATTATCTTAAATTTGGAACCTTAATTCGAGAGGAAG	1948
Db	526	TGCCCCCTTTTAGCCATGTTCCATTATCTTAAATTTGGAACCTTAATTCGAGAGGAAG	585
QY	1949	TAAGAAGGTCGTGCTGTGGCTAGGCTAGGTGAACCCCGGGTAGGGGAAAGATGTTAA	2008
Db	586	TAAGAAGGTCGTGCTGTGGCTAGGCTAGGTGAACCCCGGGTAGGGGAAAGATGTTAA	645
QY	2009	CACCTTTGACGCTTTTGGAGTTGACATGAACAGCAGGTAGTTGTTATGTAGAGCTAGTT	2068
Db	646	CACCTTTGACGCTTTTGGAGTTGACATGAACAGCAGGTAGTTGTTATGTAGAGCTAGTT	705
QY	2069	CTCAAAGCTGCCCTGCTGTTTGGAGGCGTTCACAAACAGATTGAGGCTCTTTT	2125
Db	706	CTCAAAGCTGCCCTGCTGTTTGGAGGCGTTCACAAACAGATTGAGGCTCTTTT	762

RESULT 5
BQ774398/c
LOCUS
DEFINITION UI-H-EZ1-bca-h-10-0-UI.s1 NCI_CGAP_Ch2 Homo sapiens cDNA clone

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

UI-H-EZ1-bca-h-10-0-UI 3', mRNA sequence.
BQ774398
BQ774398.1 GI:21982874
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 665)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of Orthopaedics
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA sequence: 35-59, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..665
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EZ1-bca-h-10-0-UI"
/tissue type="Chondrosarcoma Grade II"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone lib="NCI_CGAP_Ch2"
/notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCGT.
TAG_TISSUE=grade-2-chondrosarcoma
TAG_LIB=UI-H-EZ1
TAG_SEQ=ATCTAATATG"

ORIGIN

		Query Match	
		Best Local Similarity	17.2%; Score 626; DB 5; Length 665;
		Matches 626; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	2590	AACAACAACAAAAACAAGTCCAACTGATATATCTATATTTCTTAAAGT	2649
Db	665	AACAACAACAAAAACAAGTCCAACTGATATATCTATATTTCTTAAAGT	606
QY	2650	GAACGAAAGCATTTAACTGGCCAGTTTGGATTGCAATGCTGTAAGATAGAATGAAG	2709
Db	605	GAACGAAAGCATTTAACTGGCCAGTTTGGATTGCAATGCTGTAAGATAGAATGAAG	546
QY	2710	TCCTGTGAGGCTTCTCTATCTCCAGTCTATGTTTCTGGAGACCAACCATACCA	2769
Db	545	TCCTGTGAGGCTTCTCTATCTCCAGTCTATGTTTCTGGAGACCAACCATACCA	486
QY	2770	GATAATCACAAGAAAGCTTTTAAAGGCTTAAACCAAGACCTTGTCTAGATATTTT	2829

Db 485 GATAATCACAAGAAAGCTTTTAAAGGCTTAAACCAAGACCTTGTCTAGATATTTT 426

QY 2830 TAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTAAAGGGTGAG 2889

Db 425 TAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTAAAGGGTGAG 366

QY 2890 ACACAACAGTCTGACTATGAGTGAGGAAAATATCTGGTCTTTTCGTCAGTTTGGTGCA 2949

Db 365 ACACAACAGTCTGACTATGAGTGAGGAAAATATCTGGTCTTTTCGTCAGTTTGGTGCA 306

QY 2950 TTGCTGCTGCTGTTGCTACTGTTTGCCTCAACCGCTGTGTTTAAACAACGTTAAACTCTT 3009

Db 305 TTGCTGCTGCTGTTGCTACTGTTTGCCTCAACCGCTGTGTTTAAACAACGTTAAACTCTT 246

QY 3010 AGCTTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAAATTAATGATGTCTGAC 3069

Db 245 AGCTTACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAAATTAATGATGTCTGAC 186

QY 3070 ATGCTATTTTGTAGGGAGAAAATATGTGCTAATGATATTTTGAGTTAAATATCTTTTG 3129

Db 185 ATGCTATTTTGTAGGGAGAAAATATGTGCTAATGATATTTTGAGTTAAATATCTTTTG 126

QY 3130 GGGAGGATTTGCTGAAAAGTTCACATTTTGTTCACATGCTTATGCTTGGTACAGCTTAT 3189

Db 125 GGGAGGATTTGCTGAAAAGTTCACATTTTGTTCACATGCTTATGCTTGGTACAGCTTAT 66

QY 3190 GCTGCTCTTAAATTATTTTAAAAAAT 3215

Db 65 GCTGCTCTTAAATTATTTTAAAAAAT 40

RESULT 6

AW964159

LOCUS AW964159 730 bp mRNA linear EST 01-JUN-2000

DEFINITION EST376232 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.

ACCESSION AW964159

VERSION AW964159.1 GI:8153995

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 730)

AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.

TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray

JOURNAL Unpublished (2000)

COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 198

FEATURES

source

1..730

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_lib="MAGE resequences, MAGH"

/note="Vector: pBluescriptSKm"

ORIGIN

Query Match 16.4%; Score 597; DB 2; Length 730;

Best Local Similarity 100.0%; Pred. No. 1.1e-273;

Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2123 TTTTAGAATTGAATTTACTCTTCAGTATTTTCTAATGTTTCAGCTTTCTAAAGGCATATA 2182

Db 16 TTTTAGAATTGAATTTACTCTTCAGTATTTTCTAATGTTTCAGCTTTCTAAAGGCATATA 75

QY 2183 TTTTTCAAAGAAGTGAGGATGCAGTTTCTCACGTTGCAACCTATTCTGAAGTGGTTAAA 2242

Db 76 TTTTTCAAAGAAGTGAGGATGCAGTTTCTCACGTTGCAACCTATTCTGAAGTGGTTAAA 135

QY 2243 TGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAACACGGAGCTGGGCCATCGTCAGAAC 2302

Db 136 TGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAACACGGAGCTGGGCCATCGTCAGAAC 195

QY 2303 TAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAAATCATTCCTAGTACATTTTGCTAACAC 2362

Db 196 TAAGTCAGGGAAGGAGATGGATGAGAAGGCCAGAAATCATTCCTAGTACATTTTGCTAACAC 255

QY 2363 TTTATTGAGAAAATTGACCATGAATTAATGGACTCATCTTAATTTCTTCTAAGTCCATATA 2422

Db 256 TTTATTGAGAAAATTGACCATGAATTAATGGACTCATCTTAATTTCTTCTAAGTCCATATA 315

QY 2423 TAGATAGATATCTATCTGTACAGATTTCTATTTTATCCATAGATAGGTATCTATACATACA 2482

Db 316 TAGATAGATATCTATCTGTACAGATTTCTATTTTATCCATAGATAGGTATCTATACATACA 375

QY 2483 CATCTCAAGTGCATCTATTCCCACTCTCATTAATCCATCATGTTTCTTAAATTTTGTAAAT 2542

Db 376 CATCTCAAGTGCATCTATTCCCACTCTCATTAATCCATCATGTTTCTTAAATTTTGTAAAT 435

QY 2543 CTTACTGTAAAAAAGTGCACCTGAACTTCAAAAACAAAAACAAACAAACAAAA 2602

Db 436 CTTACTGTAAAAAAGTGCACCTGAACTTCAAAAACAAAAACAAACAAACAAAA 495

QY 2603 AACCAAGTCCAAACTGATATATCTCTATATTCTGTGTTAAATTCAAAAGTGAACGAAAGCATT 2662

Db 496 AACCAAGTCCAAACTGATATATCTCTATATTCTGTGTTAAATTCAAAAGTGAACGAAAGCATT 555

QY 2663 TAACGTGCCAGTTTTTGGATTGCAAAATGCTGTAAAGATATAGAAATGAAGTCTGTGAGG 2719

Db -556 TAACGTGCCAGTTTTTGGATTGCAAAATGCTGTAAAGATATAGAAATGAAGTCTGTGAGG 612

RESULT 7

CA414694/c

LOCUS CA414694 763 bp mRNA linear EST 07-NOV-2002

DEFINITION UI-H-EZO-bar-c-24-0-UI.s1 NCI_CGAP_Ch1 Homo sapiens cDNA clone

UI-H-EZO-bar-c-24-0-UI 3', mRNA sequence.

ACCESSION CA414694

VERSION CA414694.1 GI:24777345

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 763)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of Orthopedics

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA sequence: 189-213, >AT_rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

1..763

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

[illegible]

RESULT 9

RESORT 3
BQ447103/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS
COTIPOCE

SOURCE
ORGANISM

2
2
1
4
5
4
4
5

REFERENCE

AUTHORS

ATTITI

BQ447103 786 bp mRNA linear EST 29-MAY-2000
 UI-H-EU1-bad-k-13-0-UI.s1 NCI_CGAP_Ctl Homo sapiens cDNA clone
 UI-H-EU1-bad-k-13-0-UI 3', mRNA sequence.
 BQ447103
 BQ447103.1 GI:21250215
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 786)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 The following repetitive elements were found in this cDNA
 sequence: 28-52, >ATrich#Low_complexity (matched complement)
 632-675, >(CAAAA)n#Simple_repeat (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

```
Location/Qualifiers
1. .786
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="UI-H-EU1-bad-k-13-0-UI"
  /tissue_type="Osteoarthritic Cartilage"
  /dev_stage="Adult"
  /lab_host="NC10B (Life Technologies)"
  /clone_lib="NCI_CGAP_Ct1"
  /note="Organ: Knee; Vector: pT7T3-Pac (
```

TAG TISSUE=osteoarthritic cartilage

TAG_LIB=UI-H-EU1
TAG_SEQ=TGATCACGCT"

Query Match 15.4%; Score 561; DB 5; Length 786;
Best Local Similarity 99.9%; Pred. No. 1.8e-256;
Matches 681; Conservative 0; Mismatches 0; Indels

QY	2534	TTTTGTAATCTTACTGTATAAAAAAAGTGCACTGAACTTCAAAAAACAAAAACAACAA	2599
Db	713	TTTTGTAATCTTACTGTATAAAAAAAGTGCACTGAACTTCAAAAAACAAAAACAACAA	654
QY	2594	ACAACAAAAACAAGTCCAAACTGATATATCCTATATCTGTTAAAAATTCAAAAAGTGAAC	2653
Db	653	ACAACAAAAACAAGTCCAAACTGATATATCCTATATCTGTTAAAAATTCAAAAAGTGAAC	594
QY	2654	GAAAGCATTTAACTGGCCAGTTTTGATTGCAAAATGCTGTAAAGATATAGAATGAAGTCCT	2713
Db	593	GAAAGCATTTAACTGGCCAGTTTTGATTGCAAAATGCTGTAAAGATATAGAATGAAGTCCT	534
QY	2714	GTGAGGCCCTTCCTATCTCCAAGTCTATGTAATTTCTGGAGACCAACCCAGATACCAGATA	2773
Db	533	GTGAGGCCCTTCCTATCTCCAAGTCTATGTAATTTCT - GAGACCAACCCAGATACCAGATA	475
QY	2774	ATCACAAAGAAAGCTTTTTTAATAAGGCTTAAACCAAGACCTTGCTAGATATTTTTAGT	2833
Db	474	ATCACAAAGAAAGCTTTTTTAATAAGGCTTAAACCAAGACCTTGCTAGATATTTTTAGT	415
QY	2834	TTGTTGCCAAGGTAGCACCTGTGAGAAATCTCACCTTGGATGTTATGTAAAGGGGTGAGACAC	2893
Db	414	TTGTTGCCAAGGTAGCACCTGTGAGAAATCTCACCTTGGATGTTATGTAAAGGGGTGAGACAC	355
QY	2894	AACAGTCTGACTATGAGTGAGGAAATATCTGGGTCTTTTCGTCACTTTGGTGCATTTGC	2953
Db	354	AACAGTCTGACTATGAGTGAGGAAATATCTGGGTCTTTTCGTCACTTTGGTGCATTTGC	295
QY	2954	TGCTGCTGTGCTACTGTTTGCCTCAAAACGCTGTGTTTAAACAACGTTAAACTCTTAGCC	3013
Db	294	TGCTGCTGTGCTACTGTTTGCCTCAAAACGCTGTGTTTAAACAACGTTAAACTCTTAGCC	235
QY	3014	TACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATGATGCTGACATGC	3073
Db	234	TACAAGGTGGCTCTTATGTACATAGTTGTTAATACATCCAATTAATGATGCTGACATGC	175
QY	3074	TATTTTGTAGGGAGAAAAATATGTCCTAATGATATTTTGAGTTAAATATCTTTTGGGGA	3133
Db	174	TATTTTGTAGGGAGAAAAATATGTCCTAATGATATTTTGAGTTAAATATCTTTTGGGGA	115
QY	3134	GGATTTGCTGAAAAAGTTGCACCTTTTGTGTACAATGCTTATGCTTGGTACAAGCTTATGCTG	3193
Db	114	GGATTTGCTGAAAAAGTTGCACCTTTTGTGTACAATGCTTATGCTTGGTACAAGCTTATGCTG	55
QY	3194	TCCTAAATTATTTTAAAAAAAT	3215
Db	54	TCCTAAATTATTTTAAAAAAAT	33

RESULT 10
AI819798/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS
SOURCE

SOURCE
ORGANISM

NOTA

AI819798 557 bp mRNA linear EST 21-DEC-1999
wJ45a04.x1 NCI CGAP Lu19 Homo sapiens CDNA IMAGE:2405742 3'
similar to TR:035779 O35779 ENHANCER-OF-SPLIT AND HAIRY-RELATED
PROTEIN 1.; mRNA sequence.

AI819798

AI819798.1 GI:5438877

TSF

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; pr

1 (bases 1 to 557)

NCI-CGAP <http://www.nci-cgap.org>

National Cancer Instit

JOURNAL
COMMENT
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 340 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 441.
Location/Qualifiers
1. .557
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2405742"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI CGAP Lul9"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

FEATURES
source

ORIGIN

Query Match 15.3%; Score 556; DB 1; Length 557;
Best Local Similarity 100.0%; Pred. No. 4.6e-254;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 135 ATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTACTGGAACATAGAGATTTTATA 194
Db 557 ATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTACTGGAACATAGAGATTTTATA 498
QY 195 GGAATGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGAC 254
Db 497 GGAATGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGAC 438
QY 255 ACCAAGGATACCTACAAATACCGCACAGATTAATAGAAAGAAAGAAAGAGACCGAATT 314
Db 437 ACCAAGGATACCTACAAATACCGCACAGATTAATAGAAAGAAAGAAAGAGACCGAATT 378
QY 315 AATGAATGCATTGCTCAGCTGAAAGATTTACTGCTGAACATCTGAAATTGACAACTCTG 374
Db 377 AATGAATGCATTGCTCAGCTGAAAGATTTACTGCTGAACATCTGAAATTGACAACTCTG 318
QY 375 GGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACCTTAAAGCTTTAAC 434
Db 317 GGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACCTTAAAGCTTTAAC 258
QY 435 GCCTTAACCGAGCAACAGCATCAGAAGATTAATGCTTTACAGAAATGGGAGCGATCTCTG 494
Db 257 GCCTTAACCGAGCAACAGCATCAGAAGATTAATGCTTTACAGAAATGGGAGCGATCTCTG 198
QY 495 AAATCGCCCATTCAGTCGACTTGGATCGGTTCCACTCGGGATTTCAAACATGCGCCAAA 554
Db 197 AAATCGCCCATTCAGTCGACTTGGATCGGTTCCACTCGGGATTTCAAACATGCGCCAAA 138
QY 555 GAAGTCTTGCAATACCTCTCCGGTTTGGAGAGCTGGACACCCAGGAGCCCGGTGTGTC 614
Db 137 GAAGTCTTGCAATACCTCTCCGGTTTGGAGAGCTGGACACCCAGGAGCCCGGTGTGTC 78

QY 615 CAGCTGATCAACCACTTGCACGCCGTGGCCACCCAGTTCCTTGCCACCCCGCAGCTGTG 674
Db 77 CAGCTGATCAACCACTTGCACGCCGTGGCCACCCAGTTCCTTGCCACCCCGCAGCTGTG 18
QY 675 ACTCAACAGGTCCCTC 690
Db 17 ACTCAACAGGTCCCTC 2
RESULT 11
BG929136 720 bp mRNA linear EST 06-NOV-2001
LOCUS HNC59-1-B9.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
DEFINITION sequence.
ACCESSION BG929136
VERSION BG929136.1 GI:14323659
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 720)
AUTHORS Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
Sathe,G., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
Lark,M.W.
TITLE Identification and initial characterization of 5000 expressed
sequenced tags (ESTs) each from adult human normal and
osteoarthritic cartilage cDNA libraries
Osteoarthr. Cartil. 9 (7), 641-653 (2001)
JOURNAL MEDLINE
PUBMED 21482651
COMMENT 11597177
Contact: Sanjay Kumar
UW2109
GlaxoSmithKline
709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
Tel: 610-270-7245
Fax: 610-270-5598
Email: sanjay.kumar-1@gsk.com
Seq primer: T7.
Location/Qualifiers
1. .720
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="cartilage"
/lab_host="E.coli DH10 B"
/clone_lib="HNC (Human Normal Cartilage)"
/note="Vector: pSPORT I; Site_1: SalI; Site_2: NotI;
Directional"
FEATURES
source
Query Match 15.3%; Score 556; DB 4; Length 720;
Best Local Similarity 99.8%; Pred. No. 4.5e-254;
Matches 606; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
ORIGIN
QY 2587 AACAAACAACAACAAAAACAAGTCCAACTGATATATCCTATATCTGTTAAATTCAAA 2646
Db 89 AACAAACAACAACAAAAACAAGTCCAACTGATATATCCTATATCTGTTAAATTCAAA 148
QY 2647 AGTGAACGAAAGCAATTTAACTGGCCAGTTTGAATGCAATGCTGTAAAGATATAGAATG 2706
Db 149 AGTGAACGAAAGCAATTTAACTGGCCAGTTTGAATGCAATGCTGTAAAGATATAGAATG 208
QY 2707 AAGTCTCTGTGAGGCTTCCTATCTCCAAGTCTATGTTTCTGGAGACCAACAGATA 2766
Db 209 AAGTCTCTGTGAGGCTTCCTATCTCCAAGTCTATGTTTCTGGAGACCAACAGATA 268
QY 2767 CCAGATAATCACAAGAAAGCTTTTTTAAAGGCTTAAACCAAGACCTTGTCTAGATAT 2826
Db 269 CCAGATAATCACAAGAAAGCTTTTTTAAAGGCTTAAACCAAGACCTTGTCTAGATAT 328
QY 2827 TTTTAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTAAGGGT 2886

Db	329	TTTTAGTTTGTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTAAAGGGGT	388
Qy	2887	GAGACACAACAGTCTGACTATGAGTGAGGAAAATATCTGGGTCTTTTCGTCAGTTTGGTG	2946
Db	389	GAGACACAACAGTCTGACTATGAGTGAGGAAAATATCTGGGTCTTTTCGTCAGTTTGGTG	448
Qy	2947	CATTGCTGCTGCTGTTGCTACTGTTTGCCTCAAAACGCTGTGTTTAAACACAGTTAAACT	3006
Db	449	CATTGCTGCTGCTGTTGCTACTGTTTGCCTCAAAACGCTGTGTTTAAACACAGTTAAACT	508
Qy	3007	CTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAAATACATCCCAATTAAATGATGCT	3066
Db	509	CTTAGCCTACAAGGTGGCTCTTATGTACATAGTTGTTAAATACATCCCAATTAAATGATGCT	568
Qy	3067	GACATGCTATTTTGTAGGAGAAAATATGTGCTAATGATATTTTGAGTTAAAAATATCTT	3126
Db	569	GACATGCTATTTTGTAGGAGAAAATATGTGCTAATGATATTTTGAGTTAAAAATATCTT	628
Qy	3127	TTGGGGAGGATTGCTGAAAAGTTGCACCTTTTGTACAATGCTTATGCTTGGTACAAAGCT	3186
Db	629	TTGGNGAGGATTGCTGAAAAGTTGCACCTTTTGTACAATGCTTATGCTTGGTACAAAGCT	688
Qy	3187	TATGCTG 3193	
Db	689	TATGCTG 695	

RESULT 12
BG680321 630 bp mRNA linear EST 01-MAY-2001
LOCUS 602629213F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754019 5',
DEFINITION mRNA sequence.
ACCESSION BG680321
VERSION BG680321.1 GI:13911718
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 630)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10614 row: j column: 04
High quality sequence stop: 627.
Location/Qualifiers
1. .630
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754019"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source
15.2%; Score 554; DB 4; Length 630;
Best Local Similarity 100.0%; Pred. No. 4.1e-253;
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 4.1e-253;
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2123	TTTTAGAATTGAATTTACTCTTCAGTATTTTCTAATGTTTCAGCTTTCTAAAGGCATATA	2182
Db	48	TTTTAGAATTGAATTTACTCTTCAGTATTTTCTAATGTTTCAGCTTTCTAAAGGCATATA	107
Qy	2183	TTTTTCAAAGAAAGTGAGGATGAGTTCACGTTGCAACCTATTCTGAAAGTGGTTTAAA	2242
Db	108	TTTTTCAAAGAAAGTGAGGATGAGTTCACGTTGCAACCTATTCTGAAGTGGTTTAAA	167
Qy	2243	TGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAAACCGAGCTGGGCCATCGTCAGAAC	2302
Db	168	TGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAAACCGAGCTGGGCCATCGTCAGAAC	227
Qy	2303	TAACTCAGGGAAGGAGATGGATGAGAAAGGCCAGAAATCATTCCTAGTACATTTGCTAACAC	2362
Db	228	TAACTCAGGGAAGGAGATGGATGAGAAAGGCCAGAAATCATTCCTAGTACATTTGCTAACAC	287
Qy	2363	TTTATTGAGAAAATTGACCATGAATTAATGGACTCATCTTAATTTCTTCTAAGTCCATATA	2422
Db	288	TTTATTGAGAAAATTGACCATGAATTAATGGACTCATCTTAATTTCTTCTAAGTCCATATA	347
Qy	2423	TAGATAGATATCTATCTGTACAGATTTCTATTATCCATAGATAGGTATCTATACATACA	2482
Db	348	TAGATAGATATCTATCTGTACAGATTTCTATTATCCATAGATAGGTATCTATACATACA	407
Qy	2483	CATCTCAAGTGCATCTATTCCCACTCTCATTAATCCCATGTTCTCTAAATTTTGTAAAT	2542
Db	408	CATCTCAAGTGCATCTATTCCCACTCTCATTAATCCCATGTTCTCTAAATTTTGTAAAT	467
Qy	2543	CTTACTGTAAAAAAAAGTGCACCTGAACCTTCAAAAACAAAACAAAACAAAACAAA	2602
Db	468	CTTACTGTAAAAAAAAGTGCACCTGAACCTTCAAAAACAAAACAAAACAAAACAAA	527
Qy	2603	AACAAGTCCAAACTGATATATCTATATCTTCTGTTAAATTCAAAAGTGAACGAAAGCAT	2662
Db	528	AACAAGTCCAAACTGATATATCTATATCTTCTGTTAAATTCAAAAGTGAACGAAAGCAT	587
Qy	2663	TAACTGGCCAGTTT 2676	
Db	588	TAACTGGCCAGTTT 601	

RESULT 13
LOCUS CA312575/c
DEFINITION CA312575 757 bp mRNA linear EST 04-NOV-2002
UI-CF-FNO-afk-h-06-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
UI-CF-FNO-afk-h-06-0-UI 3', mRNA sequence.
ACCESSION CA312575.1 GI:24530673
VERSION CA312575
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 757)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research

Db	301	CTTGGGGCGCCGACCCCGCAGACCCCGCAGCGCCGCGACCCACCTGCCCTTCGCCGGGCC	360
QY	1515	CGCGAGCCGGGAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGGAA	1574
Db	361	CGCGAGCCGGGAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGGAA	420
QY	1575	GCTCCTGAATCCTTGCTCCCGAAGGACGGAGGTTCAAGCAGAGTGAGAAAGTTAAATA	1634
Db	421	GCTCCTGAATCCTTGCTCCCGAAGGACGGAGGTTCAAGCAGAGTGAGAAAGTTAAATA	480
QY	1635	CCCTTAAGGAGGTTCAAGCAGAGTGAGAAGTTAAAAATACCCCTT	1677
Db	481	CCCTTAAGGAGGTTCAAGCAGAGTGAGAAGTTAAAAATACCCCTT	523

RESULT 15
BF971232

LOCUS BF971232 955 bp mRNA linear EST 22-JAN-2001
DEFINITION 602273249F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4361004 5',
mRNA sequence.

ACCESSION BF971232

VERSION BF971232.1 GI:12338447

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM	Homo sapiens
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 955)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert St

Email: cgapbs-r@mail.nih.gov

Tissue procurement: ATCC

CDNA Library Preparation

CDNA Library Arranged by: The I M A G E Consortium
CDNA Library Preparation: Fire Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (FUND)
 DNA Sequencing by: Incyte Genomics Inc

DNA sequencing by: **cycle sequencing, Illumina**
Clone distribution: MGK clone distribution

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

http://image.tht.gov
http://image.tht.gov
http://image.tht.gov

Plate: LTAM10003 Row: 6 Column: 15
 Width: 200mm Height: 228

High quality sequence stop: 638.

FEATURES	
Location/Qualifiers	1. OFF

source

1. 955

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/clone="IMAGE:4361004"
```

```

/tissue_type="adrenal cortex carcinoma, cell line"

```

```
/lab_host="DH10B (phage-resistant)"
```

```
/clone_lib="NIH_MGC_84"
```

```
/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
```

NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT

primed. Average insert size 1.229 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Query Match	14.3%	Score 522;	DB 2;	Length 955;
Best Local Similarity	100.0%;	Pred. No. 8e-238;		
Matches 522; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	2566	GAACTTCAAAA CAAAAACAAAAA CAACAACAACAA CAAGTCCAAACT GTATATATCC
Db	1	GAACTTCAAAA CAAAAACAAAAA CAACAACAACAA CAAGTCCAAACT GTATATATCC

		TATATTCTGTTAAATAATCAAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTTTGATTGC	2685
QY	2626	TATATTCTGTTAAATAATCAAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTTTGATTGC	2685
Dd	61	TATATTCTGTTAAATAATCAAAAAGTGAACGAAAGCATTTAACTGGCCAGTTTTTGATTGC	120

	Qy	Db
2686	ATGCTGTAAGATATAGAAATGAAGTCCTGTGAGGCCCTCTCTATCTCCAAGTCTATGTATT	121 ATGCTGTAAGATATAGAAATGAAGTCCTGTGAGGCCCTCTCTATCTCCAAGTCTATGTATT
2745		180

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2004, 19:33:10 ; Search time 196 Seconds
(without alignments)
1420.823 Million cell updates/sec

Title: US-10-078-650-12
Perfect score: 2517
Sequence: 1 MDEGIPHLQERLLEHRDFI.....PGNPESSAQEDPSQPGKEAP 484

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2498	99.2	482	1	BHB3_HUMAN	Q9C0J9 homo sapien
2	2494	99.1	482	2	Q8TAT1	Q8tat1 homo sapien
3	2328.5	92.5	513	2	Q7YQC9	Q7yqc9 canis famil
4	1634	64.9	410	1	BHB3_MOUSE	Q99pv5 mus musculu
5	1634	64.9	410	2	BAD18970	Bad18970 mus muscu
6	1624	64.5	410	1	BHB3_RAT	Q35779 rattus norv
7	905.5	36.0	193	2	Q8CGH4	Q8cgh4 mus musculu
8	685.5	27.2	411	1	BHB2_MOUSE	Q35185 mus musculu
9	685.5	27.2	411	2	BAC33281	Bac33281 mus muscu
10	685.5	27.2	411	2	BAC41056	Bac41056 mus muscu
11	684.5	27.2	403	2	Q6NY50	Q6ny50 brachydanio
12	684.5	27.2	403	2	AAH66738	Aah66738 brachydan
13	682.5	27.1	411	1	BHB2_RAT	Q35780 rattus norv
14	679.5	27.0	411	2	Q76JQ4	Q76jq4 rattus norv
15	679.5	27.0	411	2	BAD01588	Bad01588 rattus no
16	669	26.6	412	1	BHB2_HUMAN	O14503 homo sapien
17	669	26.6	412	2	CAG33202	Cag33202 homo sapi
18	589.5	23.4	409	2	Q6GNE9	Q6gne9 xenopus lae
19	242	9.6	786	2	Q7QUD2	Q7qjd2 anopheles g
20	230.5	9.2	610	2	Q9U450	Q9u450 drosophila
21	230.5	9.2	698	2	Q9VGZ5	Q9vgz5 drosophila
22	220.5	8.8	333	2	Q8CD44	Q8cd44 mus musculu
23	220.5	8.8	339	2	Q9QUS4	Q9qus4 mus musculu
24	217.5	8.6	317	2	Q8AXV6	Q8axv6 brachydanio
25	215.5	8.6	769	2	Q9LLJ1	Q9llj1 chlamydomon
26	215	8.5	3288	2	Q7T5D9	Q7t5d9 cercopithec
27	209.5	8.3	460	2	Q9GZF7	Q9gzf7 caenorhabdi
28	207.5	8.2	3326	2	Q7T591	Q7t591 cercopithec
29	207	8.2	337	2	Q9UBP5	Q9ubp5 homo sapien
30	207	8.2	668	2	Q8MXU8	Q8mxu8 caenorhabdi
31	205.5	8.2	1953	2	Q9BIT7	Q9bit7 nephila ina

RESULT 1
BHB3_HUMAN
ID BHB3_HUMAN STANDARD; PRT; 482 AA.
AC Q9C0J9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (BHLHB3) (Differentially
DE expressed in chondrocytes protein 2) (hDEC2) (Enhancer-of-split and
DE hairy-related protein 1) (SHARP-1).
GN Name=BHLHB3; Synonyms=DEC2, SHARP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21092582; PubMed=11162494; DOI=10.1006/bbrc.2000.4133;
RA Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K.,
RA Yoshida E., Suardita K., Matsuda Y., Kato Y.;
RT "Molecular cloning and characterization of DEC2, a new member of basic
RT helix-loop-helix proteins.";
RL Biochem. Biophys. Res. Commun. 280:164-171(2001).
[2]
FUNCTION.
MEDLINE=21226716; PubMed=11278948; DOI=10.1074/jbc.M011619200;
RA Garriga-Canut M., Roopra A., Buckley N.J.;
RT "The basic helix-loop-helix protein, SHARP-1, represses transcription
RT by a histone deacetylase-dependent and histone deacetylase-independent
RT mechanism.";
RL J. Biol. Chem. 276:14821-14828(2001).
CC -!- FUNCTION: May be a transcriptional repressor that represses both
CC basal and activated transcription.
CC -!- SUBUNIT: Homodimerize.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in skeletal muscle and brain,
CC moderately expressed in pancreas and heart, weakly expressed in
CC placenta, lung, liver and kidney.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- SIMILARITY: Contains 1 orange domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

EMBL; AB044088; BAB21502.1; -.
DR PIR; JC7583; JC7583.
DR Genew; HGNC:16617; BHLHB3.
DR MIM; 606200; -.
DR GO; GO:0005634; C:nucleus; NAS.

32 204.5 8.1 290 2 Q6XLQ0 Q6xlq0 gallus gall
33 204.5 8.1 290 2 AAP44728 Aap44728 gallus ga
34 203.5 8.1 304 2 Q6XLP9 Q6xlp9 gallus gall
35 203.5 8.1 304 2 AAP44729 Aap44729 gallus ga
36 203 8.1 400 2 Q6JJ70 Q6jj70 strongyloce
37 203 8.1 400 2 AAS46235 Aas46235 strongylo
38 201.5 8.0 590 2 Q9H7S9 Q9h7s9 homo sapien
39 200.5 8.0 290 2 Q6V6Z4 Q6v6z4 branchios
40 200.5 8.0 290 2 AAQ93669 Aaq93669 branchios
41 200.5 8.0 324 2 Q9I9L0 Q9i9l0 brachydanio
42 199 7.9 328 2 Q9NQ87 Q9nq87 homo sapien
43 195.5 7.8 507 2 Q9VJ16 Q9vj16 drosophila
44 195 7.7 335 2 Q9NFW8 Q9nfw8 cupiennius
45 194 7.7 290 1 HES1_CHICK O57337 gallus gall

ALIGNMENTS

DR GO: GO:0003700; F:transcription factor activity; NAS.
DR GO: GO:0030154; P:cell differentiation; NAS.
DR GO: GO:0008283; P:cell proliferation; NAS.
DR GO: GO:0009887; P:organogenesis; NAS.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR003650; Orange.
DR Pfam: PF07527; Hairy orange; 1.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR SMART: SM00511; ORANGE; 1.
DR PROSITE: PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Repressor; Transcription regulation.
FT DNA BIND 45 57 Basic motif.
FT DOMAIN 58 100 Helix-loop-helix motif.
FT DOMAIN 129 175 Orange.
FT DOMAIN 297 431 Ala/Gly-rich.
SQ SEQUENCE 482 AA; 50497 MW; 2BECDC2FDEB7CE14 CRC64;

Query Match 99.2%; Score 2498; DB 1; Length 482;
Best Local Similarity 99.6%; Pred. No. 7.4e-122;
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MDEGPHLQERQLLEHRDFICLDYSSLYMCKPKRSMKRD DTKVSDTYKLPRLIEKKRRD 60
Db |||||
QY 1 MDEGPHLQERQLLEHRDFICLDYSSLYMCKPKRSMKRD DTK--DTYKLPRLIEKKRRD 58
Db |||||

QY 61 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTALTEQHQKIIALQNGER 120
Db |||||
QY 59 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTALTEQHQKIIALQNGER 118
Db |||||

QY 121 SLKSPIQSDLD AFHSGFQTC AKEVLQYLSRFESWTPTPREPRCVQLINHLH VAVATQFLPTPQ 180
Db |||||
QY 119 SLKSPIQSDLD AFHSGFQTC AKEVLQYLSRFESWTPTPREPRCVQLINHLH VAVATQFLPTPQ 178
Db |||||

QY 181 LLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT 240
Db |||||
QY 179 LLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT 238
Db |||||

QY 241 DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDS PAPKRMKLD SRGGSGGGPGGGAA 300
Db |||||
QY 239 DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDS PAPKRMKLD SRGGSGGGPGGGAA 298
Db |||||

QY 301 AAAAAALLGPD PAAAAALLRPDAALLSSLVAFGGGGGAPFPQ PAAAAAPFC LPPFCFLSPSA 360
Db |||||
QY 299 AAAAAALLGPD PAAAAALLRPDAALLSSLVAFGGGGGAPFPQ PAAAAAPFC LPPFCFLSPSA 358
Db |||||
QY 361 AAAYVQPF LDKSGLEKYL PAAAAAPFLLYPGIPAPAAAAA AAAAAA AAAAAA PFC LSSV 420
Db |||||
QY 359 AAAYVQPF LDKSGLEKYL PAAAAAPFLLYPGIPAPAAAAA AAAAAA AAAAAA PFC LSSV 418
Db |||||
QY 421 LSPPEKAGAAAATLLPHEVAPLGAPHQHPHGRTHLPFAGPREPGNP ESSAQEDPSQPG 480
Db |||||
QY 419 LSPPEKAGAAAATLLPHEVAPLGAPHQHPHGRTHLPFAGPREPGNP ESSAQEDPSQPG 478
Db |||||
QY 481 KEAP 484
Db |||||
QY 479 KEAP 482
Db |||||

RESULT 2
Q8TAT1 PRELIMINARY; PRT; 482 AA.
ID Q8TAT1
AC Q8TAT1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Basic helix-loop-helix domain containing, class B, 3.
GN Name=BHLHB3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025968; AAH25968.1; -.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR003650; Orange.
DR Pfam: PF07527; Hairy orange; 1.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR SMART: SM00511; ORANGE; 1.
DR PROSITE: PS50888; HLH; 1.
SQ SEQUENCE 482 AA; 50525 MW; 23EBD420C6BOCE03 CRC64;

Query Match 99.1%; Score 2494; DB 2; Length 482;
Best Local Similarity 99.4%; Pred. No. 1.2e-121;
Matches 481; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MDEGPHLQERQLLEHRDFICLDYSSLYMCKPKRSMKRD DTKVSDTYKLPRLIEKKRRD 60
Db |||||
QY 1 MDEGPHLQERQLLEHRDFICLDYSSLYMCKPKRSMKRD DTK--DTYKLPRLIEKKRRD 58
Db |||||

QY 61 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTALTEQHQKIIALQNGER 120
Db |||||
QY 59 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTALTEQHQKIIALQNGER 118
Db |||||

QY 121 SLKSPIQSDLD AFHSGFQTC AKEVLQYLSRFESWTPTPREPRCVQLINHLH VAVATQFLPTPQ 180
Db |||||
QY 119 SLKSPIQSDLD AFHSGFQTC AKEVLQYLSRFESWTPTPREPRCVQLINHLH VAVATQFLPTPQ 178
Db |||||

QY 181 LLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT 240
Db |||||
QY 179 LLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAELAAENDTDT 238
Db |||||

QY 241 DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDS PAPKRMKLD SRGGSGGGPGGGAA 300
Db |||||
QY 239 DSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDS PAPKRMKLD SRGGSGGGPGGGAV 298
Db |||||

QY 301 AAAAAALLGPD PAAAAALLRPDAALLSSLVAFGGGGGAPFPQ PAAAAAPFC LPPFCFLSPSA 360
Db |||||
QY 299 AAAAAALLGPD PAAAAALLRPDAALLSSLVAFGGGGGAPFPQ PAAAAAPFC LPPFCFLSPSA 358
Db |||||
QY 361 AAAYVQPF LDKSGLEKYL PAAAAAPFLLYPGIPAPAAAAA AAAAAA AAAAAA PFC LSSV 420
Db |||||
QY 359 AAAYVQPF LDKSGLEKYL PAAAAAPFLLYPGIPAPAAAAA AAAAAA AAAAAA PFC LSSV 418
Db |||||
QY 421 LSPPEKAGAAAATLLPHEVAPLGAPHQHPHGRTHLPFAGPREPGNP ESSAQEDPSQPG 480
Db |||||
QY 419 LSPPEKAGAAAATLLPHEVAPLGAPHQHPHGRTHLPFAGPREPGNP ESSAQEDPSQPG 478
Db |||||
QY 481 KEAP 484
Db |||||
QY 479 KEAP 482
Db |||||

Query Match		64.9%; Score 1634; DB 2; Length 410;
Best Local Similarity		72.0%; Pred. No. 3.9e-77;
Matches 350; Conservative 21; Mismatches 37; Indels 78; Gaps 16;		
QY	1	MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKSKMKRDDTKVSDTYKLEHRLIEKKRRD 60
Db	1	MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKSKMKRDDTK--DTYKLEHRLIEKKRRD 58
QY	61	RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLLKHLKALTALTEQCHQKIIALQNGER 120
Db	59	RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLLKHLKALTALTEQCHQKIIALQNGER 118
QY	121	SLKSPQSDLDFAHSGFQTCACEVLYLSRFESWTREPVCVQLINHLHVAVATQFLPTPQ 180
Db	119	SLKSPVQADLDFAHSGFQTCACEVLYLSRFESWTREPVCVQLVSHLHVAAT-----Q 172
QY	181	LLTQVPLSKGTG-APSAAGSAAAPCLERAGQKLEPLAYCVPVIQRTQPSAEIAENDTD 239
Db	173	LLTQVPSGRSGRAPCSAGAAASGPER-----VARGVPIQRTQPGTE--PEHDTD 223
QY	240	TDSGYGGEAEARPDREKKGAGASRVTIKQEPPEGDSPPAPKMKLDSRGGSGGPGGGA 299
Db	224	TDSGYGGEAE-----QGRAAVKQEPPEGDSPPAPKPKLEARG----- 260
QY	300	AAAAALLGPDPAALRRDPAALLSSLVAFGGGGGAPFPQAAAAAPFCLPFCFLSPS 359
Db	261	-----ALLGPEP-----ALLGSLVAL--GGGAPFAQP--AAAPFCLPFCFLSPS 300
QY	360	AAAYVQPFLLDKSGLEKLYLPAAAAAPFLLYPGIPAPAAAAAPFLLYPGIPAPAAAAAPFLLSPS 419
Db	301	AAAYVQPWLDKSGLDKLYLP-AAAAAPFLLYPGIP-----AAAAAPFLLSPS 350
QY	420	VLSPPPEKAGA-AAATLLPHEVAPLGAPHPQHGHGTHLPFAGPREPGNPESSAQEDPSQ 478
Db	351	VLSPPPEKAGATAGAPFLAHEVAPPGPLRPQHAHSRTHLPRA-----VNPESS-QEDATQ 404
QY	479	PGKEAP 484
Db	405	PAKDAP 410
RESULT 6		
BHB3_RAT		
ID	BHB3_RAT	STANDARD; PRT; 410 AA.
AC	O35779;	
DT	28-FEB-2003 (Rel. 41, Created)	
DT	28-FEB-2003 (Rel. 41, Last sequence update)	
DT	01-OCT-2004 (Rel. 45, Last annotation update)	
DE	Class B basic helix-loop-helix protein 3 (bHLHB3) (Enhancer-of-split and hairy-related protein 1) (SHARP-1).	
GN	Name=Bhlhb3; Synonyms=Sharp1;	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=Sprague-Dawley; TISSUE=Cerebellum;	
RX	MEDLINE=98193761; PubMed=9532582;	
RA	Rossner M.J., Doerr J., Gass P., Schwab M.H., Nave K.-A.;	
RT	"SHARPs: mammalian enhancer-of-split- and hairy-related proteins coupled to neuronal stimulation.";	
RL	Mol. Cell. Neurosci. 10:460-475(1997).	
CC	-!- FUNCTION: May be a transcriptional repressor that represses both basal and activated transcription.	
CC	-!- SUBUNIT: Homodimerize.	
CC	-!- SUBCELLULAR LOCATION: Nuclear (By similarity).	
CC	-!- TISSUE SPECIFICITY: Highly expressed in subregions of the brain, moderately expressed in skeletal muscle, heart. Weakly expressed in lung.	
CC	-!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.	
CC	-!- SIMILARITY: Contains 1 orange domain.	
CC	-!- CAUTION: Ref.1 sequence differs from that shown due to a	

14 LEHRDFIGLDYSSLY-MCKPKRSMKR-DDTKVSDTYKLPHELIEKKRRDRINECIAQLKD 71
 20 LEHGDLSCGMDFAHMYQVYKSRGIGKRSSEDSK-ETYLKLPHELIEKKRRDRINECIAQLKD 77
 72 LLPEHLKLTTLGHLEKAVVLETLTKHLKALTALTEQHQHQAIALONGERS---LKSPIQS 128
 78 LLPEHLKLTTLGHLEKAVVLETLTKHLKALTALTEQHQHQAIALONGERS---LKSPIQS 137
 129 DLDAFHSGFQCAKEVLYLSRFESWTPEPRCVQLINHLHAVATQFLPTPQLLTQVPL 188
 138 GOEMFCSGFQCAKEVLYLSRFESWTPEPRCVQLINHLHAVATQFLPTPQLLTQVPL 192
 189 SKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRT-QPSAELAENDTDTDSGYGGE 247
 193 D---SAPKAVDLKPKSPFLAKGSE-GPGKNCVPIQRTFAPSGEGSGSDTDTDSGYGGE 248
 248 AEA-----RPDREKKGAGASRVTIKQEPGGEDSPAPK-RMKLDSRGGSGGSGPG 296
 249 LEKGLRSEQPYFKSDHGRFRAVGERVSTIKQE--SEEPPTKSRMQLSEEGHFAG--- 303
 Query Match 27.2%; Score 685.5; DB 2; Length 411;
 Best Local Similarity 43.1%; Pred. No. 6e-28;
 Matches 184; Conservative 46; Mismatches 102; Indels 95; Gaps 19;

297 GGAATAAALLGPDPAATAAALLRDPDAALLSSLVAFGGGGAPFPQAPAAAAAPFCLPFCFL 356
 304 -----SDLM-----GSPFLGPHPHQPPFCLPF-YL 327
 357 SPSAAAAAYVQFLDKSGLEKLYLPAAAAAPFLLYPGIPAPAAAAAATAAATAAFAAFC 416
 328 IPPSATAYL-PM-----LEKWCYPTSV-----PVLYPGLNTSAAA----- 361
 417 LSSVLSP 423
 362 LSSFMNP 368
 RESULT 11
 Q6NY50
 ID Q6NY50 PRELIMINARY; PRT; 403 AA.
 AC Q6NY50;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Hypothetical protein zgc:76913.
 GN Name=zgc:76913;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC066738; AAH66738.1; -
 DR InterPro; IPR001092; HLH basic.
 DR InterPro; IPR003650; Orange.
 DR Pfam; PF07527; Hairy orange; 1.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00511; ORANGE; 1.
 DR PROSITE; PS50888; HLH; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 403 AA; 44682 MW; 08442DC03FFF298E CRC64;
 Query Match 27.2%; Score 684.5; DB 2; Length 403;
 Best Local Similarity 40.1%; Pred. No. 6.6e-28;
 Matches 186; Conservative 53; Mismatches 122; Indels 103; Gaps 19;
 6 PHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKR-DDTKVSDTYKLPHELIEKKRRDRINE 64

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF009330; AAB63587.1; -.
RGD; 68439; Bhlhb2.
InterPro; IPR001092; HLH_basic.
InterPro; IPR003650; Orange.
Pfam; PF07527; Hairy_orange; 1.
Pfam; PF00010; HLH; 1.
SMART; SM00353; HLH; 1.
SMART; SM00511; ORANGE; 1.
PROSITE; PS50888; HLH; 1.
DNA-binding; Nuclear protein; Transcription regulation.
DNA_BIND 53 65 Basic motif.
DOMAIN 66 108 Helix-loop-helix motif.
DOMAIN 140 184 Orange.
SEQUENCE 411 AA; 45529 MW; E56BD468D08824AD CRC64;

Query Match 27.1%; Score 682.5; DB 1; Length 411;
Best Local Similarity 43.5%; Pred. No. 8.6e-28;
Matches 186; Conservative 45; Mismatches 100; Indels 97; Gaps 20;

QY 14 LEHRDFGLDYSSLY-MCKPKRSMKR-DDTKVSDTYKLPRLHLEKRRDRINECIAQLKD 71
DB 20 LEHGDLGMDFAHMYQVYKSRGKRSSEDSK--ETYLPHRLIEKRRDRINECIAQLKD 77
QY 72 LLPEHLKLTTLGHLEKAVVLELTALKHLEKALTALTEQHQHOKIIALQN---GERSLKSPIQ 127
DB 78 LLPEHLKLTTLGHLEKAVVLELTALKHVKALTNLDQOQKIMALQSLQAGDLSGKN-IE 136
QY 128 SLDLAFHSGFQTCAKEVLYLSRPFESWTPPREPCVQLINHLHVAATQFLPTPQLLTQQVPL 187
DB 137 AGQEMFCGFGQTCAREVLQYLAKHEN--TRDLKSSQLVTHLHRVVSSELL---QGSASRKP 191
QY 188 LSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRT-QPSAELAAENDTDTDSGYGG 246
DB 192 LD---SAPKPVDFKEKPSFLAKGSE-GPGKNCVPIQRTFAPSGGEGSDTDTDSGYGG 247
QY 247 EAEA-----RPDREKKGAGASRVTIKQEPPEGDSAPK-RMKLDSRGGSGGGP 295
DB 248 ELEKGLRSEQPYFKSDHGRFTVGERVSTIKQE--SEEPPTKSRMQLSDEEGHFVG-- 303
QY 296 GGGAAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPAALAAAPFLPFCF 355
DB 304 -----SDLM-----GSPFLGPHPHQPPFCLEF-Y 326
QY 356 LSPSAAAAYVQPFLDKSGLEKLYPAAAAAPFLLYPGIPAPAAAAAFAAFAAFAAFAA 415
DB 327 LIPPSATAYL-PM-----LEKWCYPTSV-----PLLYPGLNTSAAA----- 361

RESULT 14
Q76JQ4
ID Q76JQ4 PRELIMINARY; PRT; 411 AA.
AC Q76JQ4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Enhancer of split-and hairy-related protein-2.
GN Name=SHARP-2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
TISSUE=Kidney;
PubMed=14725860;
Hirano S., Yamada K., Kawata H., Shou Z., Mizutani T., Shigematsu Y., Mayumi M., Miyamoto K.;
"The rat enhancer of split- and hairy-related protein-2 gene: hepatic expression, genomic structure, and promoter analysis.";
Arch. Biochem. Biophys. 422:81-90(2004).
EMBL; AB096137; BAD01588.1; -.
InterPro; IPR001092; HLH_basic.
InterPro; IPR003650; Orange.
Pfam; PF07527; Hairy_orange; 1.
Pfam; PF00010; HLH; 1.
SMART; SM00353; HLH; 1.
SMART; SM00511; ORANGE; 1.
PROSITE; PS50888; HLH; 1.
SEQUENCE 411 AA; 45557 MW; 2086DABB3DB82058 CRC64;

Query Match 27.0%; Score 679.5; DB 2; Length 411;
Best Local Similarity 43.1%; Pred. No. 1.2e-27;
Matches 184; Conservative 45; Mismatches 103; Indels 95; Gaps 19;

QY 14 LEHRDFGLDYSSLY-MCKPKRSMKR-DDTKVSDTYKLPRLHLEKRRDRINECIAQLKD 71
DB 20 LEHGDLGMDFAHMYQVYKSRGKRSSEDSK--ETYLPHRLIEKRRDRINECIAQLKD 77
QY 72 LLPEHLKLTTLGHLEKAVVLELTALKHLEKALTALTEQHQHOKIIALONGERS---LKSPIQS 128
DB 78 LLPEHLKLTTLGHLEKAVVLELTALKHVKALTNLDQOQKIMALQSLQAGDLSGRNIEA 137
QY 129 DLDLAFHSGFQTCAKEVLYLSRPFESWTPPREPCVQLINHLHVAATQFLPTPQLLTQQVPL 188
DB 138 GQEMFCGFGQTCAREVLQYLAKHEN--TRDLKSSQLVTHLHRVVSSELL---QGSASRKP 192
QY 189 SKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQRT-QPSAELAAENDTDTDSGYGGE 247
DB 193 D---SAPKPVDFKEKPSFLAKGSE-GPGKNCVPIQRTFAPSGGEGSDTDTDSGYGGE 248
QY 248 AEA-----RPDREKKGAGASRVTIKQEPPEGDSAPK-RMKLDSRGGSGGGPG 296
DB 249 LEKGLRSEQPYFKSDHGRFTVGERVSTIKQE--SEEPPTKSRMQLSDEEGHFVG--- 303
QY 297 GGGAAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPAALAAAPFLPFCFL 356
DB 304 -----SDLM-----GSPFLGPHPHQPPFCLEF-YL 327
QY 357 SPSAAAAYVQPFLDKSGLEKLYPAAAAAPFLLYPGIPAPAAAAAFAAFAAFAAFAA 416
DB 328 IPPSATAYL-PM-----LEKWCYPTSV-----PLLYPGLNTSAAA----- 361
QY 417 LSSVLSP 423
DB 362 LSSFMNP 368

RESULT 15
BAD01588
ID BAD01588 PRELIMINARY; PRT; 411 AA.
AC BAD01588;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Enhancer of split-and hairy-related protein-2.
GN SHARP-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
TISSUE=Kidney;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: December 15, 2004, 20:05:13 ; Search time 40 Seconds
(without alignments)
1164.223 Million cell updates/sec

Title: US-10-078-650-12
Perfect score: 484
Sequence: 1 MDEGIPHLQERQLLEHRDFI.....PGNPESQAQEDPSQPGKEAP 484

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	440	90.9	482	2 JC7583	basic helix-loop-h
2	81	16.7	410	2 JC7584	basic helix-loop-h
3	52	10.7	412	2 JC5547	basic helix-loop-h
4	16	3.3	305	2 I57039	genomic screen hom
5	16	3.3	314	2 JC5273	paired type homeob
6	16	3.3	403	2 A53662	homeotic protein H
7	16	3.3	495	1 S31223	transcription fact
8	16	3.3	513	2 A48233	polyomavirus enhan
9	16	3.3	627	2 T02610	probable YME1 ATP-
10	16	3.3	1028	2 A56038	DNA-binding protei
11	16	3.3	1213	2 S16356	ovo protein - frui
12	16	3.3	1668	2 T13748	sex comb protein -
13	15	3.1	220	2 JC5954	ribosomal protein
14	15	3.1	401	2 A48423	engrailed homeodom
15	15	3.1	420	2 T14911	bZIP DNA-binding p
16	15	3.1	443	1 I38239	transcription fact
17	15	3.1	796	2 JC7555	C14orf4 protein -
18	15	3.1	806	2 T13690	hypothetical prote
19	15	3.1	1065	2 T13230	dachshund isoform
20	15	3.1	1072	2 T13232	dachshund protein
21	15	3.1	1074	2 T13229	dachshund protein
22	15	3.1	1081	2 T13231	dachshund protein
23	15	3.1	1205	2 A55015	bumetanide-sensiti
24	15	3.1	1212	2 A57187	bumetanide-sensiti
25	15	3.1	2639	2 T31328	fibroin - Chinese
26	14	2.9	150	2 T12547	hypothetical prote
27	14	2.9	289	2 A43562	homeotic protein H
28	14	2.9	334	2 G02409	protein kinase C-b
29	14	2.9	364	2 I48188	gene NKX6.1 protei

30	14	2.9	374	2 T03875	probable homeobox
31	14	2.9	375	2 T03874	probable homeobox
32	14	2.9	494	2 A42170	zinc finger protei
33	14	2.9	552	1 WJFFEN	homeotic protein e
34	14	2.9	828	2 C88402	protein H05C05.1
35	14	2.9	1355	2 S40022	spalt protein - fr
36	14	2.9	1506	2 T51900	related to PAN2 pr
37	14	2.9	2038	2 A43742	female sterile hom
38	13	2.7	233	2 S11563	probable MASH-2 pr
39	13	2.7	238	2 A48279	achaete scute prot
40	13	2.7	323	2 S16318	homeotic protein H
41	13	2.7	331	2 B47236	zinc-finger protei
42	13	2.7	376	2 A26066	segmentation prote
43	13	2.7	378	2 A44443	basic helix-loop-h
44	13	2.7	477	2 A47236	zinc-finger protei
45	13	2.7	497	2 JC5076	myc-associated zin

ALIGNMENTS

RESULT 1
JC7583
basic helix-loop-helix protein, DEC2 - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C;Accession: JC7583
R;Fujimoto, K.; Shen, M.; Noshiro, M.; Matsubara, K.; Shingu, S.; Honda, K.; Yoshida, E.
Biochem. Biophys. Res. Commun. 280, 164-171, 2001
A;Title: Molecular cloning and characterization of DEC2, a new member of basic helix-loc
A;Reference number: JC7583; MUID:21092582; PMID:11162494
A;Accession: JC7583
A;Molecule type: mRNA
A;Residues: 1-482 <FUJ>
A;Cross-references: UNIPROT:Q9C0J9; DDBJ:AB044088
C;Comment: This protein, a novel member of the DEC subfamily of basic helix-loop-helix p
C;Genetics:
A;Gene: dec2
A;Map position: 12p11.23-p12.1
C;Keywords: transcription factor
F;1-173/Region: highly conserved #status predicted
F;130-173/Domain: Orange #status predicted <ORA>
F;286-411/Region: alanine and glycine-rich #status predicted

Query Match	90.9%;	Score 440;	DB 2;	Length 482;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 440;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	45	DTYKLPRLIEKRRDRINECIAQLKLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL	104	
Db	43	DTYKLPRLIEKRRDRINECIAQLKLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL	102	
QY	105	TEQQHQKIIALONGERSLKSPIQSDLDADFHSQFQTCAKEVLQYLSRFESWTPREPRCVQL	164	
Db	103	TEQQHQKIIALONGERSLKSPIQSDLDADFHSQFQTCAKEVLQYLSRFESWTPREPRCVQL	162	
QY	165	INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIO	224	
Db	163	INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIO	222	
QY	225	RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPFGEDSPAPKRMKL	284	
Db	223	RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPFGEDSPAPKRMKL	282	
QY	285	DSRGGSGGGPGGGAATAALLGPDPAATAALLRPDAALLSSLVAFGGGGGAPFPQAA	344	
Db	283	DSRGGSGGGPGGGAATAALLGPDPAATAALLRPDAALLSSLVAFGGGGGAPFPQAA	342	
QY	345	AAAPFCLPFCFLSPSAAAAAYVQFFLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAA	404	
Db	343	AAAPFCLPFCFLSPSAAAAAYVQFFLDKSGLEKYLYPAAAAAPFLLYPGIPAPAAAAAAA	402	
QY	405	AAAAAAAAPFCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAHPHQHGRTHLPFAGPRE	464	

Db 403 AAAAAAAAAAFCCLSSVLSPPPEKAGAAATLLPHEVAPLGAPHPQHPHGRTHLPFAGPRE 462
465 PGNPESAAQEDPSQPGKEAP 484
463 PGNPESAAQEDPSQPGKEAP 482

RESULT 2
JC7584
basic helix-loop-helix protein, DEC2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C;Accession: JC7584
R;Fujimoto, K.; Shen, M.; Noshiro, M.; Matsubara, K.; Shingu, S.; Honda, K.; Yoshida, E.
Biochem. Biophys. Res. Commun. 280, 164-171, 2001
A;Title: Molecular cloning and characterization of DEC2, a new member of basic helix-loop-helix-
A;Reference number: JC7583; MUID:21092582; PMID:11162494
A;Accession: JC7584
A;Molecule type: mRNA
A;Residues: 1-410 <FUJ>
A;Cross-references: UNIPROT:Q99PV5; DDBJ:AB044090
A;Comment: This protein, a novel member of the DEC subfamily of basic helix-loop-helix p
C;Genetics:
A;Gene: dec2
A;Map position: 6 G2-G3
C;Keywords: transcription factor
F;1-173/Region: highly conserved #status predicted
F;130-173/Domain: Orange #status predicted <ORA>
F;277-344/Region: alanine and glycine-rich #status predicted

Query Match 16.7%; Score 81; DB 2; Length 410;
Best Local Similarity 100.0%; Pred. No. 2.9e-64;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 104
DB 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHLKALTAL 102

QY 105 TEQHQKIIALQNGERSLKSP 125
DB 103 TEQHQKIIALQNGERSLKSP 123

RESULT 3
JC5547
basic helix-loop-helix factor DEC1 - human
C;Species: Homo sapiens (man)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 09-Jul-2004
C;Accession: JC5547
R;Shen, M.; Kawamoto, T.; Yan, W.; Nakamasu, K.; Tamagami, M.; Koyano, Y.; Noshiro, M.;
Biochem. Biophys. Res. Commun. 236, 294-298, 1997
A;Title: Molecular characterization of the novel basic helix-loop-helix protein DEC1 exp
A;Reference number: JC5547; MUID:97382424; PMID:9240428
A;Accession: JC5547
A;Molecule type: mRNA
A;Residues: 1-412 <SHE>
A;Cross-references: UNIPROT:O14503; DDBJ:AB004066; NID:G2308996; PIDN:BAA21720.1; PID:92
C;Comment: This protein is involved in the control of cell differentiation in several ti
F;51-108/Region: helix-loop-helix #status predicted

Query Match 10.7%; Score 52; DB 2; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.4e-38;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 TYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKH 97
DB 52 TYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKH 103

RESULT 4
I57039
genomic screen homeobox protein 2 - mouse

C;Species: Mus sp. (mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Aug-2004
C;Accession: I57039
R;Hsieh-Li, H.M.; Witte, D.P.; Saucsik, J.C.; Weinstein, M.; Li, H.; Potter, S.S.
Mech. Dev. 50, 177-186, 1995
A;Title: Gsh-2, a murine homeobox gene expressed in the developing brain.
A;Reference number: I57039; MUID:95344993; PMID:7619729
A;Accession: I57039
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-305 <RES>
A;Cross-references: GB:S79041; NID:gl042008; PIDN:AAB34947.1; PID:gl042009
C;Genetics:
A;Gene: Gsh-2
C;Superfamily: homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;204-260/Domain: homeobox homology <HOX>

Query Match 3.3%; Score 16; DB 2; Length 305;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAAAAAAAAA 413
DB 147 AAAAAAAAAAAAAAAAAAAAA 162

RESULT 5
JC5273
paired type homeobox protein, NBP - human
C;Species: Homo sapiens (man)
C;Date: 16-Apr-1997 #sequence_revision 09-May-1997 #text_change 16-Aug-2004
C;Accession: JC5273
R;Yokoyama, M.; Nishi, Y.; Yoshii, J.; Okubo, K.; Matsubara, K.
DNA Res. 3, 311-320, 1996
A;Title: Identification and cloning of neuroblastoma-specific and nerve tissue-specific
A;Reference number: JC5272; MUID:97191543; PMID:9039501
A;Contents: neuroblastoma cell
A;Accession: JC5273
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-314 <YOK>
A;Cross-references: UNIPROT:Q99453; DDBJ:D82344; NID:gl841337; PIDN:BAA11555.1; PID:dl012
C;Comment: This protein is a transcriptional repressor involved in regulating gene expres
C;Superfamily: homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;99-155/Domain: homeobox homology <HOX>

Query Match 3.3%; Score 16; DB 2; Length 314;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAAAAAAAAA 413
DB 241 AAAAAAAAAAAAAAAAAAAAA 256

RESULT 6
A53662
homeotic protein HB9 - human
C;Species: Homo sapiens (man)
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 17-Oct-1997
C;Accession: A53662
R;Harrison, K.A.; Druey, K.M.; Deguchi, Y.; Tuscano, J.M.; Kehrl, J.H.
J. Biol. Chem. 269, 19968-19975, 1994
A;Title: A novel human homeobox gene distantly related to proboscipedia is expressed in
A;Reference number: A53662; MUID:94327547; PMID:7914194
A;Accession: A53662
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-403 <HAR>
A;Cross-references: GB:U07663
A;Note: the nucleotide sequence and conceptual translation as given are self-consistent

C;Genetics:
A;Gene: GDB:HLXB9
A;Cross-references: GDB:136411; OMIM:142994
A;Map position: 1q41-1q42.1
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;244-300/Domain: homeobox homology <HOX>

Query Match 3.3%; Score 16; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAA 413
|||
Db 121 AAAAAAAAAAAAAA 136

RESULT 7
S31223
transcription factor Brn-1 - mouse
N;Alternate names: class III POU domain protein brain-1
C;Species: Mus musculus (house mouse)
C;Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S31223
R;Hara, Y.; Rovescalli, A.C.; Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 3280-3284, 1992
A;Title: Structure and evolution of four POU domain genes expressed in mouse brain.
A;Reference number: S31223; MUID:92228768; PMID:1565620
A;Accession: S31223
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-495 <HAR>
A;Cross-references: UNIPROT:P31361; EMBL:M88299; NID:G200444; PIDN:AAA39960.1; PID:G2004
C;Superfamily: transcription factor Brn-1; homeobox homology; POU domain homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;27-49/Region: glycine-rich
F;101-112/Region: alanine-rich
F;162-180/Region: histidine/proline-rich
F;186-201/Region: alanine-rich
F;236-247/Region: glycine-rich
F;267-291/Region: histidine/proline-rich
F;316-383/Domain: POU domain homology <POU>
F;402-458/Domain: homeobox homology <HOX>

Query Match 3.3%; Score 16; DB 1; Length 495;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAA 413
|||
Db 186 AAAAAAAAAAAAAA 201

RESULT 8
A48233
polyomavirus enhancer-binding protein 2 alpha chain type 1 - mouse
N;Alternate names: PEA2 alpha chain type 1; PEA2 alpha chain type 2; PEBP2 alpha chain t
C;Species: Mus musculus (house mouse)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: A48233; B48233
R;Ogawa, E.; Maruyama, M.; Kagoshima, H.; Inuzuka, M.; Lu, J.; Satake, M.; Shigesada, K.
Proc. Natl. Acad. Sci. U.S.A. 90, 6859-6863, 1993
A;Title: PEBP2/PEA2 represents a family of transcription factors homologous to the produ
A;Reference number: A48233; MUID:93342088; PMID:8341710
A;Accession: A48233
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-513 <OGA>
A;Cross-references: UNIPROT:Q08775; GB:D14636; NID:G391766; PIDN:BAA03485.1; PID:d100399
A;Accession: B48233
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-304,'L',306 <OG2>

A;Cross-references: GB:D14637; NID:G391768; PIDN:BAA03486.1; PID:G391769
C;Genetics:
A;Gene: PEBP2alphaA
C;Superfamily: transcription factor CBF alpha 2
C;Keywords: alternative splicing; DNA binding; T-cell; transcription factor; transcripti

Query Match 3.3%; Score 16; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 4.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAA 413
|||
Db 64 AAAAAAAAAAAAAA 79

RESULT 9
T02610

probable YME1 ATP-dependant proteinase [imported] - Arabidopsis thaliana
N;Alternate names: YTA11 protein homolog T19L18.5
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02610; H84656
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, August 1998
A;Description: Arabidopsis thaliana chromosome II BAC T19L18 genomic sequence.
A;Reference number: Z14681
A;Accession: T02610
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-627 <ROU>
A;Cross-references: UNIPROT:O80983; EMBL:AC004747; NID:G3413696; PID:G3413700
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84656
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-627 <STO>
A;Cross-references: GB:AE002093; NID:G3413700; PIDN:AAC31223.1; GSPDB:GN00139
C;Genetics:
A;Gene: T19L18.5; At2g26140
A;Map position: 2
A;Introns: 40/1; 119/1; 160/3; 319/3
C;Superfamily: cell division protein ftsH; FtsH/SEC18/CDC48-type ATP-binding domain homo

Query Match 3.3%; Score 16; DB 2; Length 627;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAA 413
|||
Db 598 AAAAAAAAAAAAAA 613

RESULT 10
A56038

DNA-binding protein ovo - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C;Accession: A56038
R;Garfinkel, M.D.; Wang, J.; Liang, Y.; Mahowald, A.P.
Mol. Cell. Biol. 14, 6809-6818, 1994
A;Title: Multiple products from the shavenbaby-ovo gene region of Drosophila melanogaste
A;Reference number: A56038; MUID:95021209; PMID:7935398
A;Accession: A56038
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1028 <GAR>
A;Cross-references: UNIPROT:P51521; GB:U11383; NID:G520526; PIDN:AAB60216.1; PID:G520527

This Page Blank (uspto)